俄罗斯数学
教材选译

现代几何学：方法与应用（第一卷）

几何曲面、变换群与场

（第5版）

□ Б. А. 杜布洛文　С. П. 诺维可夫　А. Т. 福明柯　著
□ 许明　译

高等教育出版社
Higher Education Press

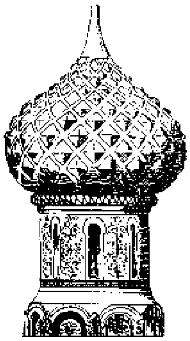
本书是莫斯科大学数学力学系对几何课程现代化改革的成果，作者之一的诺维可夫是1970年菲尔兹奖和2005年沃尔夫奖得主。全书力求以直观的和物理的视角阐述，是一本难得的现代几何方面的好书。内容包括张量分析、曲线和曲面几何、一维和高维变分法（第一卷），微分流形的拓扑和几何（第二卷），以及同调与上同调理论（第三卷）。

本书可用作数学和理论物理专业高年级和研究生的教学用书，对从事几何和拓扑研究的工作者也极有参考价值。

ISBN 7-04-018946-1

9 787040 189469 >

定价 48.10 元



俄 罗 斯 数 学

教 材 选 译

● 数学天元基金资助项目

现代几何学：方法与应用 (第一卷)

几何曲面、变换群与场

(第5版)

□ Б. А. 杜布洛文 С. П. 诺维可夫 А. Т. 福明柯 著

□ 许明 译

高等教育出版社

Higher Education Press

图字: 01–2006–3365 号

Современная геометрия: Том 1. Методы и приложения.
Геометрия поверхностей, групп преобразований и полей.
УРСС, 2001.

Originally published in Russian under the title
Modern Geometry — Methods and Applications
Part I: The Geometry of Surfaces, Transformation Groups, and Fields
Copyright © 2001 by Б. А. Дубровин, С. П. Новиков, А. Т. Фоменко
All Rights Reserved

图书在版编目 (CIP) 数据

现代几何学: 方法与应用 (第一卷) 几何曲面、变换群与场: 第 5 版 / (俄罗斯) 杜布洛文, (俄罗斯) 诺维可夫, (俄罗斯) 福明柯著; 许明译. —北京: 高等教育出版社, 2006. 9

ISBN 7-04-018946-1

I.现… II.①杜…②诺…③福…④许… III.几何学 IV.018

中国版本图书馆 CIP 数据核字 (2006) 第 064173 号

策划编辑 张小萍 **责任编辑** 蒋 青 **封面设计** 王凌波 **责任绘图** 尹文军
版式设计 史新薇 **责任校对** 金 辉 **责任印制** 朱学忠

出版发行	高等教育出版社	购书热线	010 – 58581118
社　　址	北京市西城区德外大街 4 号	免费咨询	800 – 810 – 0598
邮政编码	100011	网　　址	http://www.hep.edu.cn
总　　机	010 – 58581000		http://www.hep.com.cn
		网上订购	http://www.landraco.com
经　　销	蓝色畅想图书发行有限公司		http://www.landraco.com.cn
印　　刷	北京新丰印刷厂	畅想教育	http://www.widedu.com
开　　本	787 × 1092　1/16	版　　次	2006 年 9 月第 1 版
印　　张	23.5	印　　次	2006 年 9 月第 1 次印刷
字　　数	390 000	定　　价	48.10 元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换。
版权所有　侵权必究
物料号　18946–00

《俄罗斯数学教材选译》序

从上世纪 50 年代初起，在当时全面学习苏联的大背景下，国内的高等学校大量采用了翻译过来的苏联数学教材. 这些教材体系严密，论证严谨，有效地帮助了青年学子打好扎实的数学基础，培养了一大批优秀的数学人才. 到了 60 年代，国内开始编纂出版的大学数学教材逐步代替了原先采用的苏联教材，但还在很大程度上保留着苏联教材的影响，同时，一些苏联教材仍被广大教师和学生作为主要参考书或课外读物继续发挥着作用. 客观地说，从解放初一直到文化大革命前夕，苏联数学教材在培养我国高级专门人才中发挥了重要的作用，起了不可忽略的影响，是功不可没的.

改革开放以来，通过接触并引进在体系及风格上各有特色的欧美数学教材，大家眼界为之一新，并得到了很大的启发和教益. 但在很长一段时间中，尽管苏联的数学教学也在进行积极的探索与改革，引进却基本中断，更没有及时地进行跟踪，能看懂俄文数学教材原著的人也越来越少，事实上已造成了很大的隔膜，不能不说是一个很大的缺憾.

事情终于出现了一个转折的契机. 今年初，在由中国数学会、中国工业与应用数学学会及国家自然科学基金委员会数学天元基金联合组织的迎春茶话会上，有数学家提出，莫斯科大学为庆祝成立 250 周年计划推出一批优秀教材，建议将其中的一些数学教材组织翻译出版. 这一建议在会上得到广泛支持，并得到高等教育出版社的高度重视. 会后高等教育出版社和数学天元基金一起邀请熟悉俄罗斯数学教材情况的专家座谈讨论，大家一致认为：在当前着力引进俄罗斯的数学教材，有助于扩大视野，开拓思路，对提高数学教学质量、促进数学教材改革均十分必要.《俄罗斯数学教材选译》系列正是在这样的情况下，经数学天元基金资助，由高等教育出版社组织出版的.

经过认真选题并精心翻译校订，本系列中所列入的教材，以莫斯科大学的教材为主，也包括俄罗斯其他一些著名大学的教材. 有大学基础课程的教材，也有适合大学高年级学生及研究生使用的教学用书. 有些教材虽曾翻译出版，但经多次修订重版，面目已有较大变化，至今仍广泛采用、深受欢迎，反射出俄罗斯在出版经典教材方面所作的不懈努力，对我们也是一个有益的借鉴. 这一教材系列的出版，将中俄数学教学之间中断多年的链条重新连接起来，对推动我国数学课程设置和教学内容的改革，对提高数学素养、培养更多优秀的数学人才，可望发挥积极的作用，并起着深远的影响，无疑值得庆贺，特为之序.

李大潜

2005 年 10 月

第 2 版前言

在准备本书的第 2 版时, 作者考虑了读者的许多意见和要求: 从大学生和研究生到知名学者, 数学家和物理学家. 我们在最大范围内系统地进行了重组章节; 处理了相空间的几何理论和哈密顿系统, 并系统阐述了无穷维的 (场论方式的) 广义哈密顿系统; 另外, 作为反称张量的一个应用, 在 §18 中加进了所谓的反交换变量的积分. 系统改进的章节还包括高维变分法. 真正的扩充是从第二卷开始的, 是为了用初等的方法把读者进一步引进到流形的概念中去. 还纠正了关于刘维尔完全可积系统的证明中的某些错误, 也清除了另外一些错误以及明显的一些排版错误, 并且还扩充了文献的目录量.

作者感谢Я. Б. 泽勒罗维奇, 在我们为本书的英文和法文版作准备时他的一些意见使其中许多地方的叙述得以改进(显然, 由于这些改进才构成了现在的这个版本). 作者还要感谢本书修订版的审阅人 А. В. 波哥雷洛娃和 Ю. Г. 雷舍特尼亚克所做出的一系列有益的评注.

第 1 版前言

黎曼几何和基础拓扑到现在也还没有在大学数学教育中设置为必修课程, 就连数学系也是如此. 在标准的设置中, 替代的课程是曲线和曲面微分几何. 尽管一些地方仍在沿用这种安排, 这门课已逐渐被看成是过时的了. 那么, 这些课程究竟怎样才算是跟上了时代步伐, 就是说, 对现代数学教育而言, 现代几何的哪些内容必不可少, 同时在阐述这些内容时应把握什么样的抽象度, 对此我们并没有统一的意见.

从 1971 年开始, 莫斯科大学数学力学系力学部开始设计一门几何的现代化课程. 其内容和阐述的抽象度主要由实际应用的需要决定: 在曲线和曲面几何之外应该加进其他的专题, 即张量理论和张量的共变微分; 黎曼曲率; 测地线和变分法, 包括守恒定律和哈密顿系统; 反称张量的一种特殊情形 (即 "形式") 及其运算; 还有各种斯托克斯型的公式, 包括在各个力学分支中, 特别在流体力学和相对论等中非常有用的, 无所不在的 "广义斯托克斯公式". 许多第一流的理论物理学家和数学家一样, 认为如果还能包括有关流形、变换群、李代数以及直观拓扑学的基础概念的一些知识则是有益的. 他们还认为, 这门课程应该尽可能使用简明和具体的语言阐述, 并且只要是在实际应用的场合, 便应该用物理学家使用的术语. 以上所列举的素材组成了这个课程的最初内容, 并且在莫斯科大学以教材形式出版:

《微分几何》, 第一卷和第二卷, С. П. 诺维可夫著, 莫斯科大学力学分部, 1972.

后来, 作者修改了这个课程的许多内容, 增添了新的专题. 莫斯科大学又出版这些补充了的教材.

《微分几何》, 第三卷, С. П. 诺维可夫, А. Т. 福明柯著, 莫斯科大学力学研究所, 1974.

现在的这本书是上述这些教材经重新加工, 整理以及考虑到这些教材出版后几

何的发展而形成的. 作者认为可以把本书作为一本基础教材, 从中可以方便地提取出一门现代几何课程所必备的内容.

本书的原创概念和整个计划都来自诺维可夫. 按此计划, 原来教材中的那些素材的重新组织工作由 B. A. 杜布洛文负责, 它占了本书第一卷的多半部分; 此书的其余部分本质上是全新的. 编辑 Д. Б. 福克斯对全书的最终成形功不可没.

本书的材料大大超出了作为大学二、三年级数学教育的必修内容. 这是刻意安排的: 我们希望, 在第一卷中也应该包括一些让大学生和研究生熟悉的更为复杂的具本质性的几何概念和方法, 它涉及变换群和李代数, 场论, 变分法 (可以通过自学完成), 特别是在物理体系中起着重要作用的那些数学内容. 同时, 我们也竭力使内容的阐述和术语的使用在抽象程度上达到最小; 但是这样做常常要牺牲掉在陈述和证明上的所谓 "一般性". 事实上, 在关键的决定事物本质的那些例子中的重要的结果, 可能往往是从经典的分析与几何的初等考虑中得到的, 无需求助于现代的 "超级不变性" 的概念和记号. 但是, 对这个结果的最一般的形式表达, 特别是与其相伴的证明, 必定使抽象的形式化程度变得极为复杂. 因此, 在遇到这种场合时, 我们使用适合于这个重要例子的简洁语言进行阐述, 把一般性的证明推延在后面的章节或者干脆略而不讲. 在处理那些与现代物理学紧密相关的几何问题时, 我们曾仔细分析过物理方面的相关文献: 在量子场论方面的书籍 (例如 [36], [37]) 初始的章节中以物理学家的语言, 大篇幅地阐述了高维变分以及李群最简单表示论中的一些极重要概念; [38], [39] 则着重讲了场论的几何问题: 例如 [38] 以物理观点全面处理了黎曼几何, 其中有一些非常有用的材料. 连续介质力学和刚体理论的书 ([40], [41], [42]) 中有关张量和群论方面的许多例子也很有意思.

我们并没有打算把这套书写成 "自给自足" 的教科书: 在标准的数学教育中, 几何只是全部课程中的一个分支, 涉及分析、微分方程、代数、初等点集拓扑和测度论的各种问题, 在相关的课程中均有阐述. 我们尽力约束自己不对来自其他学科的问题进行详细讨论而只进行系统的阐述, 毕竟在标准的课程中, 它们已得到了充分的关注.

本书第二卷处理了流形的几何和拓扑, 比起第一卷, 它包含了大量的超出课程所必需的材料. 大多数关于流形的几何和拓扑的书只着眼于某个狭窄的小领域, 并且通常也十分抽象, 所用的也是一种特别适合于这个小领域的语言, 因而常常造成了不必要的复杂性. 在第二卷中我们同样尽可能忠于使阐述抽象度最小的原则; 像前面一样, 较之于一般性定理我们更优先关注有意义的例子. 另外, 我们还把章与章之间的依赖程度减到最小, 从而在内容性质所容许的范围内, 使其中每一章都可以比较容易地独立阅读. 但是我们也必须记住: 尽管可以十分容易地定义出许多拓扑概念 (例如扭结和链环、基本群、同伦群、纤维空间), 然而即便在最简单的例子中要进行非平凡的应用, 也需要发展出在经典数学中所没有的一些工具. 因此对那些虽已掌握了经典数学工具但仍不熟悉初等拓扑的读者来说, 第二卷确实比第一卷要

复杂得多. 这是无法抹平的差别. 从 20 世纪 50 年代起, 拓扑学这项工具的发展以及它对数学各分支的渗透一直在极速进行. 近年来, 拓扑方法 (有时和复代数几何结合) 实质性地应用于现代理论物理的许多问题, 可以说, 已出现了一连串的突飞猛进; 其中包括了对具有某种几何特性的特定场的量子理论, 譬如杨 – 米尔斯场和手征场; 流体晶体和超流动性理论; 广义相对论; 某些在物理上重要的非线性波动方程, 例如 KdV 方程和正弦 – 戈登方程; 某种 "长分子" 物质的统计力学也力图应用扭结和链环理论. 可惜, 我们不可能把这些应用写到本书中, 因为要把其中任一个应用处理得当就需要一段漫长的预备性旅程方能进到物理问题本身; 要是那样做, 我们就偏离初衷太远了. 但是我们在素材的选取上考虑了这些情况, 即考虑了在这些问题中所应用到的那些拓扑思想和方法, 并且还考虑到现代学派中年轻的理论物理学家们可能需要 (带有强烈目的性的) 一本可供阅读的拓扑书, 从这本书中他们可以得到想要的东西.

伴随拓扑和几何思想这二十年来的发展, 所使用的代数工具的复杂性也有了实质性的增长; 这些代数工具使用在高维的几何直观情形. 同样, 还在较深层次上使用了泛函分析、偏微分方程和复分析; 不是所有这些东西都包括在我们这本自称是初等的书里了 (事实上, 大部分这些内容也没有被包括在任何一本单独的教科书中, 只能研究各种专著和期刊).

大范围三维几何是三维空间经典微分几何的一个分支, 它是一个具直观性的, 有广泛用途的领域, 特别是凸图形理论及其应用, 它们还特别与负曲率曲面的整体问题有许多有趣的关联. 我们不是这方面的专家, 不能用十分简明和直观的例示的办法提取其精华, 从而把它们包含在一本初等的教材中. 读者可以通过 [4], [5] 和 [6] 这几本书来了解这个几何分支.

出于技术性的考虑, 讲述同调论的第三卷将由作者们另行出版.

在关于流形的拓扑和几何的所有书中, 由沙爱福和施雷发 (Seifert 和 Threlfall) 著的经典书《拓扑学》和《大范围变分法》以及更为现代的好书 [11], [12], [17], 最接近于我们在选材和处理问题时的观念了. 在创作本书的过程中, 我们以积极的态度反复思考和拓展了这些书中的素材和对它们的教学方法. 事实上, 在编写第二卷时, 我们所定的首要目标就是要写一本多少像现代版的沙爱福和施雷发的《拓扑学》, 但是涵盖的范围要更广, 尽可能运用光滑流形理论的现代方法进行重新塑造, 同时保留语言的简明, 并以新的素材加以充实. 这些新素材不但由拓扑方法的当前发展决定, 也由那些第一次接触拓扑学的读者决定; 他们希望在尽可能短的时间内学到适当的知识量. 就我们来说, 利用从物理学家所积累的方法论上的经验是完全可行的 (在第一卷中尤为如此): 通过现有的初等和熟知的手段把非初等的数学元素变得容易理解 (但是要保留数学文献的格式特点, 即主要结论的陈述均标以 "定理", "引理" 等等, 以便把它们从教材的主体中突显出来). 一般说来, 我们主张, 理解应该优先于陈述的系统化, 优先于严谨性. 有许多论断的证明细节 (姑且不论其正确与否) 在应用中

没有任何一点利用的价值. 有时在分析例子的过程中, 一些事实看起来总是合理的, 我们便在没有证明的情况下引述了这些事实并进行了应用. 对我们来说, 这似乎是合理的. 一旦对其应用完全熟悉之后, 读者最终可以在其他资料的帮助下, 自己 (如果想去做的话) 找出这些事实的证明 (为此, 我们推荐 [26]). 另外, 我们也努力把许多被省略的证明分成了能够容易解答的小片段, 放在与它们相关章节的最后面.

在第二卷的最后两章里, 我们把当代文献中许多不同类的事实组合在一起: 有关于动力体系和叶状结构的、广义相对论的以及杨 – 米尔斯场和手征场的. 在这里所阐述的思想属于不同的作者. 但是, 在一本具纯粹教科书特点的书里不可能列出一大串参考文献. 然而, 对那些将要利用研究性期刊对这些问题进行深入研究的读者来说, 一定自己会在期刊中找到合适的参考文献.

最后, 我们对莫斯科大学数学力学系的同事们表达深深的谢意: 由于他们的宝贵支持, 才使这门新的几何课程从策划到运转成为现实. 在我们系里的这些第一流的数学家中, 最应感谢的是苏维埃拓扑学派的创始人 П. С. 阿历克山得罗夫, 杰出的几何学家 П. К. 拉舍夫斯基和 Н. В. 叶非莫夫.

我们还要感谢编辑 Д. Б. 福克斯, 在文稿最后定形中他做出了巨大努力; 同样要感谢, А. Д. 阿历克山得罗夫, А. В. 波哥雷洛娃, Ю. Ф. 波里索夫, В. А. 托坡果诺夫, В. И. 库兹米诺夫, 在审阅本书时他们提供了许多有益的意见.

另外, 我们对一些学者表示特别的感谢, 是他们为本书提供了一系列非标准阐述的材料. 例如, 在普通的文献中找不到有关共形映射的刘维尔定理的证明, 是 В. А. 卓里奇与我们进行了交流. 编辑 Д. Б. 福克斯给作者指出了许多定理的简要证明. 我们还要感谢 О. И. 波哥雅夫林斯基, М. И. 莫纳斯特尔斯基, С. Г. 金吉钦, Э. Б. 温伯格, Д. В. 阿历克谢叶夫斯基, И. В. 格里伯柯夫, П. Г. 格里涅维奇.

目 录

《俄罗斯数学教材选译》序

第 2 版前言

第 1 版前言

第一章 空间区域中的几何．基本概念 1

§1. 坐标系 1

1. 空间的笛卡儿坐标(1) 2. 坐标变换(2)

§2. 欧氏空间 6

1. 欧氏空间中的曲线(6) 2. 二次型和向量(11)

§3. 黎曼和伪黎曼空间 13

1. 黎曼度量(13) 2. 闵可夫斯基度量(16)

§4. 欧氏空间的最简单的变换群 18

1. 区域的变换群(18) 2. 平面的变换(19)

3. 三维欧氏空间的运动(24) 4. 变换群的其他例子(27)

§5. 弗莱纳公式 30

1. 平面曲线(30) 2. 空间曲线, 曲率和挠率(34)

3. 依赖于参数的正交变换(37)

§6. 伪欧几里得空间 40

1. 狭义相对论的最简单概念(40) 2. 洛伦兹变换(41)

第二章　曲面论 47
§7. 空间曲面的几何 47
1. 曲面上的坐标(47)　2. 切平面(50)　3. 曲面上的度量(51)
4. 曲面的面积(54)
§8. 第二基本型 59
1. 欧氏空间中曲面上曲线的曲率(59)　2. 二次型偶对的不变量(60)
3. 第二基本型的性质(62)
§9. 球面的度量 66
§10. 在伪欧氏空间中的类空曲面 68
1. 伪球面(68)　2. $\mathbb{R}^3_1$ 中类空曲面的曲率(71)
§11. 几何中的复语言 72
1. 复坐标和实坐标(72)　2. 埃尔米特内积(73)　3. 复变换群的例子(75)
§12. 解析函数 76
1. 长度元和函数微分的复表示(76)　2. 复坐标变换(78)
3. 复空间中的曲面(81)
§13. 曲面度量的共形形式 83
1. 等温坐标、共形坐标下的高斯曲率(83)
2. 在共形形式下的球面度量和罗巴切夫斯基平面的度量(87)
3. 常曲率曲面(90)
§14. 作为 N 维空间中的曲面变换群 91
1. 在单位元的邻域内的坐标(91)　2. 矩阵的指数映射(96)　3. 四元数(99)
§15. 高维欧氏空间和伪欧氏空间的共形变换 103

第三章　张量．代数理论 110
§16. 张量的例子 110
1. 数值函数的梯度(111)　2. 黎曼度量(113)
§17. 张量的一般定义 116
1. 任意阶张量的分量的变换规律(116)　2. 张量的代数运算(121)
§18. $(0,k)$ 型张量 123
1. 表下指标张量为微分形式(123)　2. $(0,k)$ 型反称张量(125)
3. 微分形式的外积. 外代数(128)
4. $(k,0)$ 型反称张量 (多向量). 对于反交换变量的积分(128)
§19. 黎曼和伪黎曼空间中的张量 131
1. 升标和降标(131)　2. 二次型的特征值(132)　3. $*$ 算子(133)
4. 欧氏空间的张量(134)
§20. 晶体群和平面与空间旋转群的有限子群. 不变张量的例子 135

§21. 伪欧氏空间的二阶张量和它们的特征值 149
1. 反称张量. 电磁场的不变量(149)
2. 对称张量和特征值. 电磁场的能量 – 动量张量(153)
§22. 在映射下张量的行为 155
1. 具下指标的张量的一般限制运算(155) 2. 切空间的映射(157)
§23. 向量场 157
1. 微分同胚的单参数群(157) 2. 向量场的指数映射(159)
3. 李导数. 例子(160)
§24. 李代数 164
1. 李代数和向量场(164) 2. 基本的矩阵李代数(165) 3. 线性向量场(170)
4. 变换群上的左不变向量场(171) 5. 基灵度量(172)
6. 三维李代数的分类(174) 7. 共形群的李代数(175)

第四章 张量的微分学 181
§25. 反称张量的微分 181
1. 反称张量的梯度(181) 2. 形式的外微分(183)
§26. 反称张量和积分理论 189
1. 微分形式的积分(189) 2. 微分形式的例题(193)
3. 广义斯托克斯公式. 例题(198)
4. 对立方体上的广义斯托克斯定理的证明(204)
§27. 复空间中的微分形式 206
1. 算子 d' 和 d''(206) 2. 凯勒度量. 曲率形式(208)
§28. 共变微分 211
1. 欧氏联络(211) 2. 任意阶张量的共变微分(218)
§29. 共变微分和度量 221
1. 向量场的平行移动(221) 2. 测地线(223) 3. 与度量相容的联络(224)
4. 与复结构相容的联络(227)
§30. 曲率张量 230
1. 一般曲率张量(230) 2. 曲率张量的对称性. 由度量产生的曲率张量(234)
3. 例题: 基灵度量下二维和三维空间的曲率张量(235)
4. 彼得松 – 柯达齐方程. 具常负曲率的曲面和“正弦–戈登”方程(239)

第五章 变分法原理 244
§31. 一维变分问题 244
1. 欧拉 – 拉格朗日方程(244) 2. 泛函的基本例子(247)
§32. 守恒定律 250

1. 保持某个变分问题不变的变换群(250)
2. 几个例子. 守恒定律的应用(252)
§33. 哈密顿体系 261
1. 勒让德变换(261) 2. 活动坐标系(263)
3. 莫佩尔蒂和费马原理.应用(266)
§34. 相空间的几何理论 268
1. 梯度系统(268) 2. 泊松括号(271) 3. 典则变换(275)
§35. 曲面的拉格朗日函数 279
1. 轨线把和哈密顿 – 雅可比方程(279)
2. 作为动量的一阶齐次函数的哈密顿情形(282)
§36. 测地方程的二阶变分 285
1. 二阶变分公式(285) 2. 共轭点和极小性条件(288)

第六章 高维变分问题. 场及几何不变量 290
§37. 最简单的高维变分问题 290
1. 欧拉 – 拉格朗日方程(290) 2. 能量 – 动量张量(293)
3. 电磁场方程(296) 4. 引力场方程(301) 5. 皂膜(307)
6. 薄板的平衡方程(311)
§38. 拉格朗日的例子 316
§39. 广义相对论的最简单概念 319
§40. 群 $SO(3)$ 和 $O(3,1)$ 的旋量表示. 狄拉克方程和它的性质 330
1. 矩阵代数的自同构(330) 2. 群 $SO(3)$ 的旋量表示(332)
3. 洛伦兹群的旋量表示(333) 4. 狄拉克方程(336)
5. 电磁场的狄拉克方程. 电荷的共轭算子(337)
§41. 具有任意对称性的场的共变微分 338
1. 度规变换. 度规不变的拉格朗日(338) 2. 曲率形式(341)
3. 基本例子(342)
§42. 度规不变的泛函的例子. 麦克斯韦和杨 – 米尔斯方程.
具恒等于零的变分导数的泛函 (示性类) 346

参考文献 351

索引 354

第一章 空间区域中的几何. 基本概念

§1. 坐标系

我们先讨论几何学里一些基础的概念. 中学的希腊式的几何研究了简单几何图形的各种度量性质. 其中要解决的基本问题是找出在三角形和其他多边形中长度和角之间的关系, 除此之外, 在这些基础之上计算了曲面面积和某些立体的体积. 中学几何的中心概念是直线线段 (或圆弧) 的长, 两条相交线 (直的或弯曲的) 之间的夹角.

整个解析 (坐标) 几何的主要目的是在平面或三维空间的笛卡儿坐标系下描述用公式给出的几何图形. 与 "中学的" 几何相比, 唯一的差别在于方法上的不同, 而研究的对象是一样的. 微分几何也处在同样的状态: 相同的主题, 但在这里更加精妙地使用了微分学和线性代数的手段. 这样, 在引进了对一般的光滑图形的研究后, 微分几何便扩充了它研究对象的范围.

1. 空间的笛卡儿坐标

我们基本的几何思想如下:

(1) 几何是在某个空间中展开的, 而此空间由点 $P, Q, \cdots$ 组成.

(2) 像在解析几何那样, 在此空间可以引进笛卡儿坐标, 这表示对空间每个点建立一组相对应的实数 $x^1, x^2, \cdots, x^n$, 称为它的坐标, 它们满足下面两个性质:

a) 不同的点对应不同的坐标组; 这意味着, 具有坐标 $(x^1, \cdots, x^n), (y^1, \cdots, y^n)$ 的两个点 P, Q 相重合当且仅当 $x^i = y^i, i = 1, 2, \cdots, n$ 成立.

b) 反之, 每组 $(x^1, \cdots, x^n)$, 其中 x^i 为任意实数, 应对应于空间中的某一个点.

定义 1.1. 称一个引进了上述性质的笛卡儿坐标 $(x^1, \cdots, x^n)$ 的空间为n 维笛

卡儿空间,[①]记为 $\mathbb{R}^n$. 称数 n 为此空间的维数.

我们常常就称数组 $(x^1,\cdots,x^n)$ 本身为这个空间的点. 笛卡儿空间的最简单例子是数直线, 它是一维的笛卡儿空间. 这里的点只有一个坐标 $x^1(n=1)$, 另一个例子是解析几何中的, 这是在平面上的 (二维笛卡儿空间) 和 "通常的" 即三维的空间中的笛卡儿坐标 (图 1), 这些笛卡儿空间对于解决中学几何的问题是完全适合的.

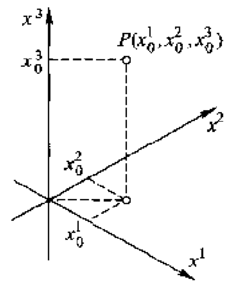


图 1

我们给出一个不太熟悉的然而极其重要的笛卡儿空间的例子. 现代物理学告诉我们, 时间和空间不是完全分离的, 不相重叠的概念, 而是出现在四维的时空连续体中. 对这些现象的自然顺序的数学阐述的证明是极方便的.

我们的时空连续体的点被当作为事件. 对每个事件我们派定一个实数的有序四元组 (t,x^1,x^2,x^3), 其中 t 是事件发生的 "瞬时", 而 x^1,x^2,x^3 为事件的 "空间位置" 的坐标. 数 (t,x^1,x^2,x^3) 便是时空连续体的笛卡儿坐标.因此, 时空连续体是个四维的笛卡儿空间. 现在我们可以不再理会把坐标 (t,x^1,x^2,x^3) 当作事件的时间和空间的解释. 经典几何的三维空间只不过是由 $t=$ 常数定义的水平曲面. 一个物体 (在每个瞬间可以抽象地看成一个点 (所谓的质点)) 的路径在时空空间中则是四维空间中一条曲线段 (或弧)$x^\alpha(t),\alpha=1,2,3$, $t_1\leqslant t\leqslant t_2$. 我们称此曲线为这个点的世界线 (图 2). 我们也会考虑三维甚至二维时空连续体, 它们分别由三联组 (t,x^1,x^2) 和偶对 (t,x^1) 为坐标, 这是因为对这些空间而言可以比较容易画出清晰的图像.

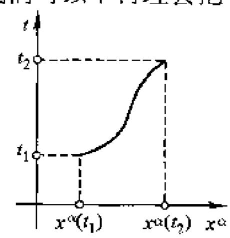


图 2

2. 坐标变换

假设在一个 n 维笛卡儿空间中我们有一个实函数 $f(P)$, 即对此空间中每个点 P 指派一个实数的函数. 由于空间中每个点以其 n 个坐标值出现, 故我们可以把 f 想成是 n 个实变量的函数: 如果 $P=(x^1,\cdots,x^n)$, 则 $f(P)=f(x^1,\cdots,x^n)$. 我们将只涉及连续的 (甚至为连续可微) 函数 $f(x^1,\cdots,x^n)$, 我们所考虑的函数常常不是对空间 $\mathbb{R}^n$ 中每个点都有定义, 而仅仅定义在它的部分或更精准地说在它的 "区域" 上.

定义 1.2. 一个区域或一个无边界区域(用其他术语也称为 "开集") 是 $\mathbb{R}^n$ 中的一个点集 D, 使得对它中的每个点也包含了充分靠近此点的所有点.

更准确地说, 对区域 D 中每个点 $P_0=(x_0^1,\cdots,x_0^n)$, 存在数 $\varepsilon>0$ 使得满足不等式

$$|x^i-x_0^i|<\varepsilon,\quad i=1,\cdots,n$$

的所有点 $P=(x^1,\cdots,x^n)$ 也在 D 中.

① 可能这不是个通用的术语, 希望不要使读者困惑.

定义 1.3. 一个带边界的区域是由无边界区域加进所有边界点得到. 区域的边界由所有边界点组成.

无边界区域的最简单的例子是整个空间 $\mathbb{R}^n$. 另一个简单的例子是平面中满足 $x_1^2+x_2^2<\rho^2$ 的点 (x_1,x_2) 的集合 (半径 $\rho>0$ 的开圆盘). 对应的带边界的区域由满足 $x_1^2+x_2^2\leqslant\rho^2$ 的点 (x_1,x_2) 组成. 这个例子是一种具有典型意义的定义. 有下面简单的定理.

定理 1.1. 设在空间 $\mathbb{R}^n$ 给出了一组连续函数 $f_1(P),\cdots,f_m(P),P=(x^1,\cdots,x^n)$. 考虑满足不等式

$$f_1(P)<0,f_2(P)<0,\cdots,f_m(P)<0$$

的点 P 的集合 D. 于是 D 是无边界区域.

证明 设 $P_0=(x_0^1,\cdots,x_0^n)$ 在 D 中, 即 $f_1(P_0)<0,\cdots,f_m(P_0)<0$. 由关于连续函数保持符号的定理得到: 对每个 j 可以找到这样的 $\varepsilon_j>0$, 使得由不等式 $|x^i-x_0^i|<\varepsilon_j,i=1,2,\cdots,n$ 得到 $f_j(P)<0$. 选取 $\varepsilon=\min(\varepsilon_1,\cdots,\varepsilon_n)$, 我们便看出 D 包含了所有满足 $|x^i-x_0^i|<0$ 的点 $(x^1,\cdots,x^n)$. 因此 D 是无边界区域. □

注 沿曲线在区域 $D:f_j(P)<0,j=1,\cdots,m$ 中的运动, 由于 f_j 的连续性, 只可能达到那些满足 $f_j(P)\leqslant0$ 的点 P. 如果函数 $f_1,\cdots,f_m$ 满足某些简单的解析条件 (在第二卷中将详细阐述), 则对任意 j 使 $f_j(P)=0$ 的任意点 P 都可以达到. 在下面遇到的所有例子都满足这些限制条件. 因此, 在这些条件下, 不等式 $f_j(P)\leqslant0,j=1,\cdots,m$ 的解给出了一个带边界的区域.

空间中有界区域的概念是非常重要和常常遇到的, 即离坐标原点足够远的点都不属于它的那种区域.

在整个空间 $\mathbb{R}^n$ 中的坐标 $(x^1,\cdots,x^n)$ 显然给出了任意区域 D 中点的坐标, 但它们没有取所有的值. 对于定义在区域中的函数 $f(x^1,\cdots,x^n)$ 可以像对定义在整个空间 $\mathbb{R}^n$ 中的那样谈及连续性和可微性.

假设在同一区域中给出任意其他的坐标 $(z^1,\cdots,z^n)$. 我们可以写成

$$\begin{aligned}x^i&=x^i(z^1,\cdots,z^n),\quad i=1,2,\cdots,n\\ z^j&=z^j(x^1,\cdots,x^n),\quad j=1,2,\cdots,n.\end{aligned}\tag{1}$$

上述等式只不过表明可以把笛卡儿坐标系 $(x^1,\cdots,x^n)$ 与新的坐标系 $(z^1,\cdots,z^n)$ 进行比较, 因此可以通过新的坐标来表示原来的坐标, 反之亦然.

我们首先考察空间中的线性坐标变换:

$$x^i=\sum_{j=1}^n a_j^i z^j,\quad i=1,\cdots,n\tag{2}$$

(较简短地可写为 $x^i=a_j^iz^j$, 其中表明按重复的指标取和). 由线性代数知道, 通过 x 能够表出 z 的充分必要条件是矩阵 $A=(a_j^i)$ 具有逆矩阵 $B=A^{-1}=(b_j^i)$. 逆矩阵由

方程 $b^i_j a^j_k = \delta^i_k$ 定义, 其中

$$\delta^i_k = \begin{cases} 1 & i = k, \\ 0 & i \neq k \end{cases}$$

(克罗内克符号), 并表明为对 j 求和. 于是, 点 P 的笛卡儿坐标 $x^1, \cdots, x^n$ 借助于矩阵 $A = (a^i_j)$ 通过新的数组 $z^1, \cdots, z^n$ 表出. 等式 (2) 可以简单地写为

$$X = AZ, \quad X = (x^1, \cdots, x^n), \quad Z = (z^1, \cdots, z^n).$$

等式 (2) 意味着, 如果点 P 对应了坐标 $x^1, \cdots, x^n$, 于是在新坐标系下它对应于数组 $z^1, \cdots, z^n$, 使得 $x^j = a^j_i z^i$. 我们看到矩阵 A 应该是可逆的, 即具有非零的行列式 (非异). 这时新的坐标可以通过老的坐标表出

$$Z = BX, \quad z^j = b^j_k x^k \tag{3}$$

(按 k 求和).

现在考虑任意的新坐标 $x^i = x^i(z^1, \cdots, z^n), i = 1, \cdots, n$, 其中的函数 $x^i(z^1, \cdots, z^n)$ 假定为连续可微的 (即光滑).

我们假定新的坐标代表了所讨论区域中每一个点, 就是说, 在所讨论的区域中任意数组 $(x^1_0, \cdots, x^n_0)$ 至少对应于一个数组 $(z^1_0, \cdots, z^n_0)$ 使得 $x^i_0 = x^i(z^1_0, \cdots, z^n_0), i = 1, \cdots, n$.

定义 1.4. 称点 $P = (x^1_0, \cdots, x^n_0)$ 为坐标系$(z^1, \cdots, z^n)$ 的非异点是说, 如果矩阵

$$A = (a^i_j) = \left(\frac{\partial x^i}{\partial z^j}\right)\bigg|_{z^1 = z^1_0, \cdots, z^n = z^n_0}, \tag{4}$$

具有非零行列式 (非退化), 其中 $z^1_0, \cdots, z^n_0$ 满足 $x^i(z^1_0, \cdots, z^n_0) = x^i_0, i = 1, \cdots, n$.

称矩阵 A 为所给变换的雅可比矩阵, 表示为 $\hat{J} = \left(\dfrac{\partial x}{\partial z}\right)$. 雅可比矩阵的行列式被称做雅可比, 以 J 表示

$$J = \det\left(\frac{\partial x}{\partial z}\right) = \det \hat{J}.$$

由数学分析课程知道下面的关于逆变换的定理 (一般隐函数定理的特殊情形):

如果已知到新坐标的变换 $x^i = x^i(z), i = 1, \cdots, n$, 其中 $x^i_0 = x^i(z^1_0, \cdots, z^n_0)$, 在 $z^1 = z^1_0, \cdots, z^n = z^n_0, J = \det\left(\dfrac{\partial x}{\partial z}\right) \neq 0$, 则在点 $(x^1_0, \cdots, x^n_0)$ 的充分小邻域内可以通过 $x^1, \cdots, x^n$ 表出坐标 $z^1, \cdots, z^n$, 使得 $z^i = z^i(x)$, 其中 $z^i_0 = z^i(x^1_0, \cdots, x^n_0) i = 1, \cdots, n$. 此时矩阵 $(b^i_j) = \left(\dfrac{\partial z^i}{\partial x^j}\right)$, 即逆变换的雅可比矩阵, 它是矩阵 $(a^k_l) = \left(\dfrac{\partial x^k}{\partial z^l}\right)$ 的逆矩阵, 满足

$$\frac{\partial z^i}{\partial x^j}\frac{\partial x^j}{\partial z^k} = \delta^i_k \tag{5}$$

(回想一下: 这表示按 j 求和).

对 $n=1$ 而言, 这个陈述可表达为: 如果 $x=x(z), x(z_0)=x_0$ 且在 $z=z_0, \frac{dx}{dz}\neq 0$, 则可以通过 x 表示出 z 使得 $z_0=z(x_0)$, 且点 x_0 的充分小的邻域内有 $\frac{dz}{dx}\frac{dx}{dz}=1$.

在上述情形的分析中如果设坐标变换是线性的, 即 $X=AZ$, 或 $x^i=a^i_jz^j$, 这时的雅可比矩阵 $\left(\frac{\partial x}{\partial z}\right)$ 与矩阵 A 相同, 有 $a^i_k=\frac{\partial x^i}{\partial z^k}$ 为常数. 那么, 如果 $\det A\neq 0$ 则在整个空间上的逆变换为 $Z=BX$, 其中 B 为 A 的逆矩阵.

我们考察其他的由解析几何中取来的关于平面和空间坐标的例子.

(1) 考虑平面上的极坐标 r, φ, 其中

$$x^1=r\cos\ \varphi,\quad x^2=r\sin\ \varphi. \tag{6}$$

这里的 $r\geqslant 0$.

另外, 偶对 (r,φ) 和 $(r,\varphi+2k\pi)$ 在 k 为整数时代表了同一个点 $P=(x^1,x^2)$. 因而如果要求 $0\leqslant\varphi<2\pi$, 则 φ 为唯一确定的坐标. 在 $r=0$ 需要进一步说明, 这时所有的偶对 $(0,\varphi)$ 表示了同一个点 (坐标原点), 因此在坐标原点会发生一些很不好的情形. 我们已知原点就是极坐标系的奇点. 构造雅可比矩阵如下

$$A=\begin{pmatrix}\frac{\partial x^1}{\partial r} & \frac{\partial x^1}{\partial\varphi}\\ \frac{\partial x^2}{\partial r} & \frac{\partial x^2}{\partial\varphi}\end{pmatrix}=\begin{pmatrix}\cos\ \varphi & -r\sin\varphi\\ \sin\ \varphi & r\cos\varphi\end{pmatrix}. \tag{7}$$

我们有雅可比

$$J=\det A=r\geqslant 0.$$

因此雅可比只在 $r=0$ 时为 0. 如果以 x^1, x^2 表示 r, 则 $r=\sqrt{(x^1)^2+(x^2)^2}$. 这个函数在 $x^1=0, x^2=0$ 处不可微. 在区域 $\{(r,\varphi)|r>0, 0<\varphi<2\pi\}$ 中, 新坐标系是完全唯一确定的, 没有奇点.

(2) 具笛卡儿坐标 x^1,x^2,x^3 的三维笛卡儿空间中的柱坐标 $z^1=r, z^2=\varphi, z^3=z$, 其中

$$x^1=r\cos\ \varphi,\quad x^2=r\sin\ \varphi,\quad x^3=z. \tag{8}$$

在这里 $r=0$ 给出了直线, 即 z 轴, 沿着它坐标系 "变坏" 了. 事实上, 雅可比矩阵在这里的形式为

$$A=\begin{pmatrix}\cos\ \varphi & -r\sin\ \varphi & 0\\ \sin\ \varphi & r\cos\ \varphi & 0\\ 0 & 0 & 1\end{pmatrix}, \tag{9}$$

它的雅可比只在 $r=0$ 时为 0. 在区域 $r>0$ 中此坐标系没有奇点. 像上面一样, 在 $0<\varphi<2\pi$ 区域内坐标 φ 是唯一确定的.

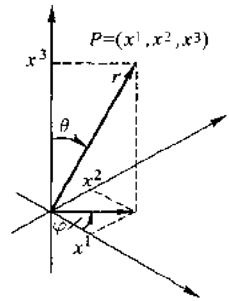


图 3

(3) 三维空间中的球面坐标 $z^1=r, z^2=\theta, z^3=\varphi$ (图 3). 我们有

$$x^1=r\cos\varphi\sin\theta,\quad x^2=r\sin\varphi\sin\theta,\quad x^3=r\cos\theta,$$

$$r\geqslant 0,\quad 0\leqslant\theta\leqslant\pi,\quad 0\leqslant\varphi<2\pi. \tag{10}$$

对于球面坐标, 其雅可比矩阵具有下面的形式

$$A=\begin{pmatrix}\cos\varphi\sin\theta & r\cos\varphi\cos\theta & -r\sin\varphi\sin\theta\\ \sin\varphi\sin\theta & r\sin\varphi\cos\theta & r\cos\varphi\sin\theta\\ \cos\theta & -r\sin\theta & 0\end{pmatrix}. \tag{11}$$

雅可比 $J=\det A$ 有形式

$$J=r^2\sin\theta.$$

我们看出, 在区域 $r>0, \theta\neq 0, \pi$ 中此雅可比不为 0.

球面坐标系在区域 $r>0, 0<\theta<\pi, 0<\varphi<2\pi$ 中是唯一确定的, 并且没有奇点. 点 $r=0$(θ,φ 任意) 和 $\theta=0,\pi$(r,φ 任意) 是球面坐标系的奇点.

§2. 欧氏空间

除了前面一节所罗列的概念外, 作为我们几何思想的基础是空间中线段长度的概念和两条相交线在交点度量角的概念. 我们的这些思想的基础在于我们 (多少有些近似地) 生活在三维的欧氏空间里, 在其中存在具特殊性质的笛卡儿坐标.

1. 欧氏空间中的曲线

假设在三维空间中有笛卡儿坐标, 使得如果三维空间中点 P 坐标为 (x^1,x^2,x^3), 点 Q 坐标为 (y^1,y^2,y^3), 则连接点 P 和 Q 的直线段的长度平方等于 $l^2=(x^1-y^1)^2+(x^2-y^2)^2+(x^3-y^3)^2$. 称这样的空间为欧几里得空间, 而具这种性质的笛卡儿坐标为欧几里得坐标, 简称欧氏坐标.

由线性代数知识知道, 把向量与欧氏空间中点联系起来有其方便之处, 我们有坐标原点即点 O, 称由 O 起始到点 P 终止的向量为点 P 的径向量. 点 P 的坐标 (x^1,x^2,x^3) 称做向量坐标. 由点 O 分别引向点 P 和 Q 的两个向量 $\xi=(x^1,x^2,x^3)$, $\eta=(y^1,y^2,y^3)$ 可以按坐标相加, 得到坐标为 $(x^1+y^1,x^2+y^2,x^3+y^3)$ 的向量 $\xi+\eta$, 也可以用数去乘向量. 具坐标 $e_1=(1,0,0), e_2=(0,1,0), e_3=(0,0,1)$ 的单位向量 e_1,e_2,e_3 的长度为 1. 后面将会看到, 它们是两两相互垂直的. 任意一个有坐标

(x^1, x^2, x^3) 的向量 ξ 有形式 $\xi = x^1e_1 + x^2e_2 + x^3e_3$. 这里所有空间是三维的, 即 $n = 3$. 对于任意 n 的情形, 自然是完全类似的. 因此欧氏空间常常被看作线性 (或向量) 空间, 其中点 (径向量终点)$\xi = (x^1, \cdots, x^n)$ 和 $\eta = (y^1, \cdots, y^n)$ 之间的距离平方由 $l^2 = \sum_{i=1}^{n}(x^i - y^i)^2$ 给出. 在三维空间中我们有 $n = 3$, 平面则是 $n = 2$, 而 $n > 3$ 则是它们的直接推广.

在欧氏空间中有一个被称做向量内积的重要运算.

定义 2.1. 如果已知向量 $\xi = (x^1, \cdots, x^n), \eta = (y^1, \cdots, y^n)$, 则称数.

$$\langle \xi, \eta \rangle = \sum_{i=1}^{n} x^i y^i \tag{1}$$

为它们的欧几里得内积.

这个内积具有下列重要性质:

a) $\langle \xi, \eta \rangle = \langle \eta, \xi \rangle$;

b) $\langle \lambda_1\xi_1 + \lambda_2\xi_2, \eta \rangle = \lambda_1\langle \xi_1, \eta \rangle + \lambda_2\langle \xi_2, \eta \rangle$, 其中 λ_1 和 λ_2 为任意实数;

c) $\langle \xi, \xi \rangle > 0$, 其中 $\xi \neq 0$.

具有形如 (1) 的内积的笛卡儿坐标 $x^1, \cdots, x^n$ 被称做欧几里得坐标.

利用内积的概念可以重新表述, 从其径向量 $\xi = (x^1, \cdots, x^n)$ 的点到其径向量 $\eta = (y^1, \cdots, y^n)$ 的点的线段长的平方等于向量 $\xi - \eta$ 自身的内积, 而任意向量 $\zeta = (z^1, \cdots, z^n)$ 的长度等于 $\sqrt{\langle \zeta, \zeta \rangle}$. 常常以 $|\zeta| = \sqrt{\langle \zeta, \zeta \rangle}$ 表示向量 ζ 的长度. 性质 c) 表明所有非零向量具有正的长度.

从解析几何知道, 对于两个向量 $\xi = (x^1, \cdots, x^n), \eta = (y^1, \cdots, y^n)$ 之间的夹角也可以通过内积来表出

$$\cos\varphi = \frac{\langle \xi, \eta \rangle}{\sqrt{\langle \xi, \xi \rangle \langle \eta, \eta \rangle}} = \frac{\langle \xi, \eta \rangle}{|\xi||\eta|}, 0 \leqslant \varphi \leqslant \pi. \tag{2}$$

因此, 长度与角和向量间的内积概念紧密相关. 更进一步说, 内积将被当作基础的首要概念, 几何正是建于其上的.

现在设在欧氏空间中有曲线, 它由参数形式给出

$$x^1 = f^1(t), \cdots, x^n = f^n(t), \tag{3}$$

其中参数 t 经过由 a 到 b 的区间, 而 $f^i(t)$ 是参数 t 的光滑函数, 其中 $i = 1, \cdots, n$. 称向量

$$v(t) = \left(\frac{df^1}{dt}, \cdots, \frac{df^n}{dt}\right) \tag{4}$$

为曲线在时刻 t 的切向量或速度向量.

定义 2.2. 称数

$$l = \int_a^b \sqrt{\langle v(t), v(t)\rangle} dt = \int_a^b |v(t)| dt \tag{5}$$

为曲线的长.

换句话说, 称速度向量的长度的积分为曲线的长度. ①

如果曲线 $x^i = f^i(t), i = 1, \cdots, n$ 和另一条曲线 $x^i = g^i(t), i = 1, \cdots, n$ 在 $t = t_0$ 相交 (即 $f^i(t_0) = g^i(t_0), i = 1, \cdots, n$), 于是可以论及在交点处这些曲线间的夹角. 我们通过

$$v = \left(\frac{df^1}{dt}, \cdots, \frac{df^n}{dt}\right)\Big|_{t=t_0}, \quad w = \left(\frac{dg^1}{dt}, \cdots, \frac{dg^n}{dt}\right)\Big|_{t=t_0} \tag{6}$$

表示在 $t = t_0$ 处曲线相应的切向量.

定义 2.3. 称满足

$$\cos\ \varphi = \frac{\langle v, w\rangle}{|v||w|} \tag{7}$$

的角 $\varphi(0 \leqslant \varphi \leqslant \pi)$ 为这两条曲线在它们于 $t = t_0$ 相交的交点处的曲线间夹角.

上述最后的两个定义应该是在数学分析课程中证明了的重要事实. 然而事实上它们可以被视为原始的定义. 仅仅需要验证这些定义与在欧氏空间中曲线间的夹角及曲线长度的直观概念相一致. 从中学的几何课程知道, 半径为 R 的圆周长为 $2\pi R$. 又从解析几何知道, 线段或具坐标 $(y^1, \cdots, y^n)$ 的向量 ξ 的长等于 $\sqrt{(y^1)^2 + \cdots + (y^n)^2}$ (由毕达哥拉斯定理). 我们来对线段和圆周进行计算, 以确信我们关于长度的定义给出同样的结果.

1) 线段. 为简单起见, 设其由坐标原点为起点. 于是它由公式 $x^i = y^i t, i = 1, \cdots, n, 0 \leqslant t \leqslant 1$ 给出. 当 $t = 0$ 时所有的坐标 $x^i = 0$, 而当 $t = 1$ 时所有坐标 $x^i = y^i$, 即在向量 ξ 的终点. 我们的直线段的长度等于

$$l = \int_0^1 \sqrt{\left(\frac{dx^1}{dt}\right)^2 + \cdots + \left(\frac{dx^n}{dt}\right)^2}\, dt = \sqrt{(y^1)^2 + \cdots + (y^n)^2},$$

于是我们得到了所熟知的关于线段长的公式.

2) 圆周. 它由 (在平面上的) 公式 $x^1 = R\cos\ t, x^2 = R\sin\ t$, 其中 $0 \leqslant t \leqslant 2\pi$ 给出. 这里的速度向量的形式为 $v(t) = (-R\sin\ t, R\cos\ t)$, 于是长度等于

$$l = \int_0^{2\pi} \sqrt{R^2\sin^2 t + R^2\cos^2 t}\, dt = 2\pi R. \tag{8}$$

对于圆周我们同样得到了所要的答案.

另外, 我们的长度定义也满足这样的要求, 即由两段构成的曲线的长等于这些段的长度之和.

①作者不想要求对长度概念作公理化的处理. 与其从一组公理导出此定义的唯一性, 还不如就把此定义当作公理.

但是, 我们注意到关于长度的公式 (5) 依赖于参数化的曲线 $x^i = f^i(t), i = 1, \cdots, n, a \leqslant t \leqslant b$. 简单地说, 沿着曲线随参数 t 一起在 a 与 b 之间以速度 $v(t) = \left(\frac{df^1}{dt}, \cdots, \frac{df^n}{dt}\right)$ 运动, 而这个沿曲线运动的速度 $v(t)$ 显式地出现在公式

$$l = \int_a^b |v(t)| dt \tag{9}$$

中.

如果我们沿同一曲线以不同的速度运动会怎样呢? 从点 $P = (f^1(a), \cdots, f^n(a))$ 到点 $Q = (f^1(b), \cdots, f^n(b))$ 前进. 如果从点 P 到点 Q 走了同样长度但所具有的速度不一样, 我们会得到同一个数吗?

这个问题的准确提法是这样的: 假设给出一个在 a' 到 b' 变化的新参数 $\tau(a' \leqslant \tau \leqslant b')$, 而参数 t 以 τ 的函数形式表出, $t = t(\tau), t(a') = a, t(b') = b$, 且 $\frac{dt}{d\tau} > 0$. 后面这个不等式表明我们的参数 τ 在曲线上以参数 t 同样的方式移动. 于是我们的曲线表示为

$$x^i = f^i(t) = f^i(t(\tau)) = g^i(\tau), \quad i = 1, \cdots, n. \tag{10}$$

按参数 τ 移动的速度为

$$w(\tau) = \left(\frac{dg^1}{d\tau}, \cdots, \frac{dg^n}{d\tau}\right), \quad a' \leqslant \tau \leqslant b'. \tag{11}$$

在新参数下曲线段的长等于

$$l' = \int_{a'}^{b'} |w(\tau)| d\tau. \tag{12}$$

我们来证明

$$l' = \int_{a'}^{b'} |w(\tau)| d\tau = l = \int_a^b |v(t)| dt.$$

计算出向量 $w(\tau)$ 的长:

$$\begin{aligned} |w(\tau)| &= \sqrt{\sum_{i=1}^n \left(\frac{dg^i}{d\tau}\right)^2} = \sqrt{\sum_{i=1}^n \left(\frac{df^i}{dt}\frac{dt}{d\tau}\right)^2} \\ &= \left|\frac{dt}{d\tau}\right| \sqrt{\sum_{i=1}^n \left(\frac{df^i}{dt}\right)^2} = \frac{dt}{d\tau}|v(t)|, \end{aligned}$$

后一个等式是由于 $\frac{dt}{d\tau} > 0$. 因此

$$\begin{aligned} l' &= \int_{a'}^{b'} |w(\tau)| d\tau = \int_{a'}^{b'} |v(t(\tau))| \frac{dt}{d\tau} d\tau \\ &= \int_a^b |v(t)| dt. \end{aligned}$$

这就是所要证明的.

结论 曲线段的长不依赖于经过这条曲线段的速度.

因此, 我们的长度定义满足了对于这个量的直观观念所关切的所有必要的要求.

例 2.1. 假设在平面上给出函数 $x^2 = f(x^1)$ 图像的曲线. 可以简单地取 x^1 为参数 $t: x^1 = t, x^2 = f(t)$. 速度向量具形式 $v = \left(1, \frac{df}{dx^1}\right)$ 而曲线长为

$$l = \int_a^b |v|dt = \int_a^b \sqrt{1 + \left(\frac{df}{dx^1}\right)^2} dx^1. \tag{13}$$

在任意光滑曲线 (速度参数 v 不为零的那种曲线) 上可以选 l(长度) 为参数, 使得速度向量为单位向量:$|v(l)| = 1$. 称这样的参数 l 为自然参数. 对其有 $\int_a^b |v|dl = b - a$. 因此, 自然参数具有简单的几何性质: 它等于所通过的曲线段的长.

现在假定在具欧氏坐标 $(x^1, \cdots, x^n)$ 的欧氏空间中给出了另一个坐标系 $(z^1, \cdots, z^n)$, 使得 $x^i = x^i(z^1, \cdots, z^n), i = 1, \cdots, n$. 设在新坐标下以参数给出了曲线 $z^i = z^i(t), i = 1, \cdots, n$. 这时在原来的欧氏坐标中同一条曲线具有形式

$$x^j - x^j(z(t)) = h^j(t), \quad j = 1, \cdots, n.$$

我们称向量 $v_z(t) = (v_z^1, \cdots, v_z^n)$, 其中

$$v_z^j = \frac{dz^j}{dt}, \quad j - 1, \cdots, n \tag{14}$$

为在坐标$(z^1, \cdots, z^n)$下的曲线的速度向量. 在原来的坐标 $(x^1, \cdots, x^n)$ 下速度向量 $v = v_x = \left(\frac{dh^1}{dt}, \cdots, \frac{dh^n}{dt}\right)$. 这是个向量, 它在点 $P = (h^1(t), \cdots, h^n(t))$ 取值; 它与 v_z 是同一个向量, 但在点 $P = (z^1(t), \cdots, z^n(t))$ 取值. 点 P 是同一个, 向量也是同一个, 但在两个不同的坐标系 (z) 和 (x) 下表出. 我们要弄清楚在坐标变换下速度的坐标是如何变化的. 我们有

$$v_x^i = \frac{dh^i}{dt} = \frac{\partial x^i}{\partial z^j}\frac{dz^j}{dt} = \frac{\partial x^i}{\partial z^j}v_z^j \tag{15}$$

(记住: 按 j 求和). 速度向量的平方具有形式

$$|v_x|^2 = \sum_{i=1}^n \left(\frac{dh^i}{dt}\right)^2 = \sum_{i=1}^n \left(\frac{\partial x^i}{\partial z^j}v_z^j\right)^2 = g_{jk}v_z^j v_z^k, \tag{16}$$

其中

$$g_{jk} = \sum_{i=1}^n \frac{\partial x^i}{\partial z^j}\frac{\partial x^i}{\partial z^k}. \tag{17}$$

结论 在任意坐标 $(z^1, \cdots, z^n), x = x(z)$ 下，曲线速度向量 $v_z = \left(\frac{dz^1}{dt}, \cdots, \frac{dz^n}{dt}\right)$ 的自身内积由公式

$$|v_z|^2 = |v_x|^2 = g_{jk}\frac{dz^j}{dt}\frac{dz^k}{dt} \tag{18}$$

给出，其中

$$g_{jk} = \sum_{i=1}^{n}\frac{\partial x^i}{\partial z^j}\frac{\partial x^i}{\partial z^k}.$$

2. 二次型和向量

我们在前面已看到，曲线的速度向量的坐标在坐标变换 $x = x(z)$ 下变为

$$v_x^i = \frac{\partial x^i}{\partial z^j}v_z^j \tag{19}$$

或缩写为

$$v_x = Av_z,$$

其中 $A = \left(\frac{\partial x}{\partial z}\right)$，即在前一节所定义的变换的雅可比矩阵. 我们在这里并没有用到坐标 $(x^1, \cdots, x^n)$ 是欧几里得的这个性质. 变换规则 (19) 可以当作向量的一般定义的一个基本性质.

定义 2.4. 称数组 $(\xi^1, \cdots, \xi^n)$ 为关于坐标系 $(x^1, \cdots, x^n)$ 在点 $P = (x_0^1, \cdots, x_0^n)$ 的一个向量. 并且如果两个坐标系 $(x^1, \cdots, x^n)$ 和 $(z^1, \cdots, z^n)$ 由变换 $x = x(z)$ 相联系，同时 $x^i(z_0^1, \cdots, z_0^n) = x_0^i, i = 1, \cdots, n$, 那么对于新的坐标系 $(z^1, \cdots, z^n)$，同一个向量在点 $(z_0^1, \cdots, z_0^n)$ 则由另一组数 $(\zeta^1, \cdots, \zeta^n)$ 给出，它们之间由基本公式

$$\xi^i = \left.\frac{\partial x^i}{\partial z^j}\right|_{z^k = z_0^k}\zeta^j \tag{20}$$

相联系.

应该注意到，在这个向量定义中主要是变换规律 (20) 的形式.

我们考虑另一个常常遇见的几何对象：函数的梯度. 我们习惯说的是在笛卡儿坐标 $(x^1, \cdots, x^n)$ 下数值函数 $f(x^1, \cdots, x^n)$ 的梯度 (例如在 $n = 3$ 的情形)，这是个向量，其分量表示为

$$\text{grad} f = \left(\frac{\partial f}{\partial x^1}, \cdots, \frac{\partial f}{\partial x^n}\right). \tag{21}$$

令 $\xi_j = \frac{\partial f}{\partial x^j}, j = 1, \cdots, n$. 我们考虑同一个函数在另一组坐标 $z^1, \cdots, z^n$ 下其梯度是如何表示的，其中 $x = x(z)$. 我们有

$$\text{grad} f(x^1(z), \cdots, x^n(z)) = \left(\frac{\partial f}{\partial z^1}, \cdots, \frac{\partial f}{\partial z^n}\right),$$

$$\frac{\partial f}{\partial z^i}=\frac{\partial f}{\partial x^j}\frac{\partial x^j}{\partial z^i}, i=1,\cdots,n.$$

以 η_i 表示分量 $\dfrac{\partial f}{\partial z^i}$, 在新坐标系下这个梯度的分量为

$$\eta_i=\frac{\partial x^j}{\partial z^i}\xi_j. \tag{22}$$

我们看出在坐标变换下函数的梯度改变与向量的改变不同. 在后面我们将这个量称为余向量.

现在设有欧氏坐标系 $(x^1,\cdots,x^n)$, 而 $\xi_1=(\xi_1^1,\cdots,\xi_1^n)$ 和 $\xi_2=(\xi_2^1,\cdots,\xi_2^n)$ 为两个从点 $P=(x_0^1,\cdots,x_0^n)$ 出发的向量. 在坐标系 $(z^1,\cdots,z^n)$ 下, 设 $x=x(z),x(z_0)=x_0$, 这两个向量具有相应的坐标 $(\eta_1^1,\cdots,\eta_1^n)$ 和 $(\eta_2^1,\cdots,\eta_2^n)$, 它们以前述的坐标公式相联系.

$$\xi_1^i=a_j^i\eta_1^j,\quad \xi_2^i=a_j^i\eta_2^j,$$

其中 (a_j^i) 为在 $z^k=z_0^k,k=1,\cdots,n$ 处的雅可比矩阵. ξ_1 和 ξ_2 在原来坐标系下的内积有形式

$$\langle\xi_1,\xi_2\rangle=\sum_{i=1}^n\xi_1^i\xi_2^i=\delta_{ij}\xi_1^i\xi_2^j. \tag{23}$$

在新坐标系下它等于

$$\langle\xi_1,\xi_2\rangle=\sum_{i=1}^n(a_j^i\eta_1^j)(a_k^i\eta_2^k)=g_{jk}\eta_1^j\eta_2^k, \tag{24}$$

其中矩阵 $G=(g_{jk})$,

$$g_{jk}=\sum_{i=1}^n a_j^i a_k^i=\delta_{sq}a_j^s a_k^q, \tag{25}$$

这就是出现在前一小节中, 在解决关于在任意坐标下计算曲线长度的问题时的同一矩阵. 因此在新坐标中向量的内积决定了同一个矩阵 $G=(g_{ij})$. 以代数的语言来说, 公式 (25) 表示

$$G=A^{\mathrm{T}}A, \tag{26}$$

其中 "T" 代表矩阵的转置. 我们来解释矩阵 G 的分量 g_{ij} 在转换到新坐标时是如何变化的. 设在同一区域中给定了新坐标 $y^1,\cdots,y^n$, 而 $z^j=z^j(y^1,\cdots,y^n),j=1,\cdots,n$. 令 $B=(b_j^i)=\left(\dfrac{\partial z^i}{\partial y^j}\right)$. 我们知道, 当在坐标 $y^1,\cdots,y^n$ 下, 向量 ξ_1,ξ_2 具有分量 $(\zeta_1^1,\cdots,\zeta_1^n),(\zeta_2^1,\cdots,\zeta_2^n)$, 其中

$$\eta_1^i=b_j^i\zeta_1^j,\quad \eta_2^i=b_j^i\zeta_2^j. \tag{27}$$

假设在坐标 (y) 下对内积表达式所给出的矩阵为 (h_{ij}). 这表明

$$\langle \xi_1, \xi_2 \rangle = h_{kl}\zeta_1^k\zeta_2^l = g_{ij}\eta_1^i\eta_2^j. \tag{28}$$

利用等式 (27) 我们得到

$$h_{kl}\zeta_1^k\zeta_2^l = (b_k^i g_{ij} b_l^j)(\zeta_1^k\zeta_2^l). \tag{29}$$

由此有

$$h_{kl} = b_k^i g_{ij} b_l^j. \tag{30}$$

于是, $H = B^{\mathrm{T}}GB$.

定义 2.5. 称关于坐标系 $(x^1, \cdots, x^n)$ 的数组 $g_{ij}, i, j = 1, \cdots, n, g_{ij} = g_{ji}$ 为在点 $x_0^1, \cdots, x_0^n$ 的 (在向量上的)二次型, 并且如果两个坐标系 $(x^1, \cdots, x^n)$ 和 $(z^1, \cdots, z^n)$ 有变换 $x = x(z)$, 满足 $x^i(z_0^1, \cdots, z_0^n) = x_0^i, i = 1, \cdots, n$, 那么对于新的坐标系 $(z^1, \cdots, z^n)$, 这个二次型给出了数组 $h_{kl}, k, l = 1, \cdots, n, h_{kl} = h_{lk}$, 它们与原来的数组由公式

$$g_{ij} = \left.\frac{\partial x^k}{\partial z^j}\right|_{z^\alpha = z_0^\alpha} h_{kl} \left.\frac{\partial x^l}{\partial z^i}\right|_{z^\beta = z_0^\beta} \tag{31}$$

相关联. 用矩阵形式表示即

$$G = A^{\mathrm{T}}HA.$$

如果在点 P 给出了二次型 g_{ij}, 在坐标变换下按规律 (31) 变化, 那么在点 P 的切向量上可以定义二次 (或双线性) 函数 $\{\xi, \xi\}$ (或 $\{\xi, \eta\}$), 即令

$$\{\xi, \xi\} = g_{ij}\xi^i\xi^j \ (\{\xi, \eta\} = g_{ij}\xi^i\eta^j).$$

由变换规律 (31) 知道, 这样定义的函数不依赖于坐标系的选取而只依赖于点 P 和向量 ξ (或向量 ξ 和 η).

§3. 黎曼和伪黎曼空间

1. 黎曼度量

我们已经讨论了在空间或空间区域中的长度或者称之为度量的概念. 在具坐标 $(x^1, \cdots, x^n)$ 的 n 维空间中光滑曲线 $x^i = x^i(t)$ 的长, 按定义由公式 (2.5) ①

$$l = \int_a^b |\dot{x}(t)| dt, \quad \dot{x} = v = \frac{dx}{dt}, \tag{1}$$

并且要求在空间中每一点上的曲线速度向量的长度概念都已事先定义好了.

①这表示参看 §2 的公式 (5). 在后面均采用这种参照记号.

黎曼度量被认为是在所给坐标系中对向量 $\xi = (\xi^1, \cdots, \xi^n)$ 的长由下面公式给出

$$|\xi|^2 = g_{ij}\xi^i\xi^j. \tag{2}$$

这表明 $|\xi|^2$ 是在上一节所说意义下的向量 ξ 的二次函数. 向量长度应不依赖于坐标系的选取, 因此 g_{ij} 在坐标变换下应该像二次型的分量那样变化, 即按公式 (2.31) 那样. 由此, 我们引进黎曼度量的概念.

定义 3.1. 称在每点有定义的并且光滑依赖于这些点的切向量的一个正定二次型为这个空间的某区域上的一个黎曼度量.

利用在上一节中所给出的二次型的定义可以把黎曼度量的定义转换为下面的形式:

定义 3.2. 称在具任意坐标 $(z^1, \cdots, z^n)$ 的空间某区域中的函数组 $g_{ij} = g_{ji}(z^1, \cdots, z^n), i, j = 1, \cdots, n$, 其中矩阵 (g_{ij}) 为正定, 为这个区域上的一个黎曼度量.

如果在这同一区域上给出了一个新的坐标 $(y^1, \cdots, y^n)$ 且 $z^i = z^i(y^1, \cdots, y^n), i = 1, \cdots, n$, 那么在新坐标中这个黎曼度量由函数组 $g'_{ij} = g'_{ji}(y^1, \cdots, y^n), i, j = 1, \cdots, n$ 定义, 其中

$$g'_{ij} = \frac{\partial z^k}{\partial y^i} g_{kl} \frac{\partial z^l}{\partial y^j}. \tag{3}$$

矩阵 (g_{ij}) 的正定性表示 $g_{ij}\xi^i\xi^j > 0$, 其中的向量 ξ 不为零.

如果给出了黎曼度量, 则曲线 $z^i = z^i(t)$ 的长度等于

$$l = \int_a^b \sqrt{g_{ij}(z(t)) \frac{dz^i}{dt} \frac{dz^j}{dt}} dt. \tag{4}$$

定义 3.3. 设 $\xi = (\xi^1, \cdots, \xi^n)$ 和 $\eta = (\eta^1, \cdots, \eta^n)$ 为在点 $P = (z_0^1, \cdots, z_0^n)$ 的两个向量. 称数 $\langle \xi, \eta \rangle$ 为它们的内积, 其中

$$\langle \xi, \eta \rangle = g_{ij}(z_0^1, \cdots, z_0^n)\xi^i\eta^j. \tag{5}$$

(g_{ij}) 和向量坐标 (ξ^i) 的变换规律 (3) 和 (2.20) 保证了从同一个点出发的两个向量的内积与坐标系的选取无关. 从不同点出发的两个向量的内积在坐标变换下不是个不变量.

如果已知有两条曲线 $z^i = f^i(t), z^i = h^i(t)$, 它们相交于 $t = t_0$. 称数 φ $(0 \leqslant \varphi \leqslant \pi)$ 为它们之间的夹角, 其中

$$\cos\varphi = \frac{\langle \xi, \eta \rangle}{|\xi||\eta|};$$

这里的 $\langle \xi, \eta \rangle = g_{ij}\xi^i\eta^j, |\xi| = \sqrt{\langle \xi, \xi \rangle}$, 而 ξ, η 为在交点的两个速度向量. (我们注意到, φ 的实数性质来自柯西 – 布涅可夫斯基不等式 $|\langle \xi, \eta \rangle| \leqslant |\xi||\eta|$ (由黎曼度量的正定性出发证明它).)

例 3.1. 欧几里得度量.

a) $n=2$. 在欧氏坐标 $x^1=x, x^2=y$

$$g_{ij}=\delta_{ij}=\begin{cases}1 & i=j;\\ 0 & i\neq j;\end{cases}\quad (g_{ij})=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

在极坐标 r,φ 下

$$x^1=r\cos\ \varphi,\ x^2=r\sin\ \varphi\ (\text{参看§1}),$$
$$r=z^1,\ \varphi=z^2,\ (g'_{ij})=\begin{pmatrix}1 & 0\\ 0 & r^2\end{pmatrix}.$$

这表明对曲线 $r=r(t),\varphi=\varphi(t),a\leqslant t\leqslant b$,

$$l=\int_a^b\sqrt{\left(\frac{dr}{dt}\right)^2+r^2\left(\frac{d\varphi}{dt}\right)^2}dt.$$

b) $n=3$. 在欧氏坐标 x^1,x^2,x^3 下我们有 $g_{ij}=\delta_{ij}$. 在柱坐标中 ($y^1=r,y^2=\varphi,y^3=z$, 参看 §1)

$$(g'_{ij})=\begin{pmatrix}1 & 0 & 0\\ 0 & r^2 & 0\\ 0 & 0 & 1\end{pmatrix},\quad l=\int_a^b\sqrt{\left(\frac{dr}{dt}\right)^2+r^2\left(\frac{d\varphi}{dt}\right)^2+\left(\frac{dz}{dt}\right)^2}dt.$$

在球面坐标中 ($y^1=r,y^2=\theta,y^3=\varphi$, 参看 §1)

$$(g'_{ij})=\begin{pmatrix}1 & 0 & 0\\ 0 & r^2 & 0\\ 0 & 0 & r^2\sin^2\theta\end{pmatrix},$$

$$l=\int_a^b\sqrt{\left(\frac{dr}{dt}\right)^2+r^2\left(\frac{d\theta}{dt}\right)^2+r^2\sin^2\theta\left(\frac{d\varphi}{dt}\right)^2}dt.$$

度量常常写成微分 dl 的平方的形式

$$dl^2=g_{ij}dz^idz^j. \tag{6}$$

(严格地说, dz^i 定义为 $dz^i=\dot{z}^idt$ (对 dl 是相似的), 其中 $z^i=z^i(t)$ 为所讨论的曲线.)

对于考察过的例子我们有:

$$\left.\begin{array}{ll}\text{对笛卡儿坐标} & dl^2=\sum_{i=1}^n(dx^i)^2,\\ \text{对极坐标} & dl^2=(dr)^2+r^2(d\varphi)^2,\\ \text{对柱坐标} & dl^2=(dr)^2+r^2(d\varphi)^2+(dz)^2,\\ \text{对球面坐标} & dl^2=(dr)^2+r^2[(d\theta)^2+\sin^2\theta(d\varphi)^2].\end{array}\right\} \tag{7}$$

定义 3.4. 称矩阵 $g_{ij}=g_{ji}(z)$ 为欧几里得的是说如果能找到坐标 $x^1,\cdots,x^n$, $x^i=x^i(z)$, 满足

$$\det\left(\frac{\partial x^i}{\partial z^j}\right)\neq 0,\quad g_{ij}=\sum_{k=1}^{n}\frac{\partial x^k}{\partial z^i}\cdot\frac{\partial x^k}{\partial z^j}.$$

这时在坐标 $x^1,\cdots,x^n$ 下

$$g'_{ij}=\delta_{ij}=\begin{cases}1, & i=j,\\ 0, & i\neq j.\end{cases}$$

称坐标 $x^1,\cdots,x^n$ 为欧几里得坐标.

上面所考虑的所有度量的例子都是在不同坐标下的欧氏坐标. 下一章我们将考虑其他的黎曼度量的例子.

2. 闵可夫斯基度量

假设数 $g_{ij}=g_{ji}(z),i,j=1,\cdots,n$ 使得矩阵 (g_{ij}) 非异, 即 $\det(g_{ij})\neq 0$, 并且形式 $g_{ij}\xi^i\xi^j$ 是 (不定的) 非正的, 这时我们说它是一个伪黎曼度量.

如果 g_{ij} 为伪度量, $g^0_{ij}=g_{ij}(z^1_0,\cdots,z^n_0)$ 为它在某个点 $P=(z^1_0,\cdots,z^n_0)$ 上的值, 那么由线性代数知道, 变换 $\xi^i=\lambda^i_k\eta^k$ 可以把二次型 $g^0_{ij}\xi^i\xi^j$ 转化为

$$\eta_1^2+\cdots+\eta_p^2-\eta_{p+1}^2-\cdots-\eta_n^2,$$

其中 p 和 q $(p+q=n)$ 分别为二次型 $g^0_{ij}\xi^i\xi^j$ 的正和负的指标. 不难看出 (参看节尾的习题) 数 p 和 q 被完全决定, 即惯性指标不依赖于坐标系. 由于连续性, 二次型 $g_{ij}\xi^i\xi^j$ 的惯性指标在所考虑的点 P 的邻域内所有的点上取同一个 (整数) 值 (但是一般来说, 在此点邻域内每点同时化成同一个标准形式却是不可能的); 在这种情形下, 我们称 g_{ij} 为 (p,q)型的伪黎曼矩阵.

定义 3.5. 称矩阵 $g_{ij}=g_{ji}(z)$ 为伪欧几里得矩阵, 如果存在新的坐标 $x^1,\cdots,x^n,x^i=x^i(z),\det\left(\frac{\partial x^i}{\partial z^j}\right)\neq 0$, 使得

$$g_{ij}=\frac{\partial x^1}{\partial z^i}\frac{\partial x^1}{\partial z^j}+\cdots+\frac{\partial x^p}{\partial z^i}\frac{\partial x^p}{\partial z^j}-\frac{\partial x^{p+1}}{\partial z^i}\frac{\partial x^{p+1}}{\partial z^j}-\cdots-\frac{\partial x^n}{\partial z^i}\frac{\partial x^n}{\partial z^j}.$$

在此新坐标系下

$$g'_{ij}=0,\quad 当\ i\neq j,$$
$$g'_{ii}=1,\quad 当\ i\leqslant p,$$
$$g'_{ii}=-1,\quad 当\ i\geqslant p+1.$$

称坐标 $(x^1,\cdots,x^n)$ 为 (p,q)型伪欧几里得坐标, 其中 $q=n-p$. 在空间 $\mathbb{R}^n$ 中定义向量 $\xi=(\xi^1,\cdots,\xi^n)$ 和 $\eta=(\eta^1,\cdots,\eta^n)$ 的 "内积" 为

$$\langle\xi,\eta\rangle_{p,q}=\xi^1\eta^1+\cdots+\xi^p\eta^p-\xi^{p+1}\eta^{p+1}-\cdots-\xi^n\eta^n,\tag{8}$$

则可以引进 (p,q) 型的伪度量; 这时通常的坐标 $x^1,\cdots,x^n$ 便是伪欧几里得的了. 具有这样度量的空间 $\mathbb{R}^n$ 也被称做伪欧几里得的, 并以 $\mathbb{R}^n_{p,q}$ 表示.

可以认为 $p\leqslant[n/2]$, 如有必要可经变换 $g_{ij}\to -g_{ij}$ 化为这种形式.

空间 $\mathbb{R}^4_{1,3}$ 的情形特别重要. 这是狭义相对论的空间 (闵可夫斯基空间). 在狭义相对论中假定了在 §1 中定义的时空连续体是闵可夫斯基空间 $\mathbb{R}^4_{1,3}$. 回想一下, 在时空连续体中的点被给予了笛卡儿坐标 (t,x^1,x^2,x^3), 这里的第一个坐标 t 是时间的维数, 而坐标 (x^1,x^2,x^3) 表示长度的维度. 相应的伪欧几里得坐标便是 $x^0=ct,x^1,x^2,x^3$, 其中 c 是具有速度量纲 (长度/时间) 的常数, 即在真空中的光速.

长度元的平方 dl^2 具有形式

$$dl^2=(dx^0)^2-(dx^1)^2-(dx^2)^2-(dx^3)^2. \tag{9}$$

如果有两个点 (或事件)$P_1=(x_1^0,x_1^1,x_1^2,x_1^3), P_2=(x_2^0,x_2^1,x_2^2,x_2^3)$, 则我们称量

$$|P_1-P_2|^2=(x_1^0-x_2^0)^2-(x_1^1-x_2^1)^2-(x_1^2-x_2^2)^2-(x_1^3-x_2^3)^2 \tag{10}$$

为事件 P_1 和 P_2 之间的时空间隔. 量 $|P_1-P_2|^2$ 可以为正, 为负, 也可 (对不同的点 P_1 和 P_2) 为零 (参见 §6).

在这一节最后我们考虑在空间 $\mathbb{R}^3_{1,2}$ 中一个有关坐标的有用的例子, 即伪球面坐标. 假设在 $\mathbb{R}^3_{1,2}$ 中的伪欧氏坐标为 x^0,x^1,x^2. 定义伪球面坐标 ρ,χ,φ 为

$$\left.\begin{aligned} x^0&=\rho\,\mathrm{ch}\,\chi, \\ x^1&=\rho\,\mathrm{sh}\,\chi\cos\varphi, \\ x^2&=\rho\,\mathrm{sh}\,\chi\sin\varphi, \end{aligned}\right\}\quad \begin{aligned} &-\infty<\rho<\infty, \\ &0\leqslant\chi<\infty, \\ &0\leqslant\varphi<2\pi. \end{aligned} \tag{11}$$

因此,

$$(x^0)^2-(x^1)^2-(x^2)^2=\rho^2\geqslant 0.$$

于是, 坐标 ρ,χ,φ 只在区域 $(x^0)^2-(x^1)^2-(x^2)^2>0$ 中有定义. 在空间 $\mathbb{R}^3_{1,2}$ 中这个区域是圆锥 $(x^0)^2=(x^1)^2+(x^2)^2$ 的内部 (图 4), 这个区域中所有点 (除去 x^0 轴上的点) 都是伪球面坐标的非异点. 长度元的平方 dl^2 在此区域中具形式

$$dl^2=d\rho^2-\rho^2[(d\chi)^2+\mathrm{sh}^2\chi(d\varphi)^2]. \tag{12}$$

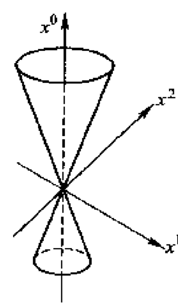


图 4

不难在锥的外部也引进伪球面坐标, 它们的公式是

$$\left.\begin{aligned} x^0&=\rho\,\mathrm{sh}\,\chi, \\ x^1&=\rho\,\mathrm{ch}\,\chi\cos\varphi, \\ x^2&=\rho\,\mathrm{ch}\,\chi\sin\varphi, \end{aligned}\right\}\rho>0. \tag{13}$$

这个情形在应用上不太重要.

习题

3.1. 证明伪黎曼度量的类型不依赖于坐标的选择.

§4. 欧氏空间的最简单的变换群

1. 区域的变换群

假设在 n 维空间中有两个区域: 区域 Ω_x, 其坐标为 $x^1, \cdots, x^n$ 和区域 Ω_z, 坐标为 $z^1, \cdots, z^n$. 另外假设区域 Ω_z 中每个点被指定了区域 Ω_x 中一个对应的点, 使 $x^i = x^i(z^1, \cdots, z^n), i = 1, \cdots, n$. 如果坐标可以反过来通过 $x^1, \cdots, x^n$ 表示 $z^1, \cdots, z^n$, 即 $z^j = z^j(x^1, \cdots, x^n), j = 1, \cdots, n$, 那么就说给出了从区域 Ω_z 到区域 Ω_x 的一个变换. 当然, 这时我们要求函数 $x^i(z^1, \cdots, z^n)$ 和它们的反函数 $z^j(x^1, \cdots, x^n)$ 是光滑的. 于是雅可比 $\det\left(\dfrac{dx}{dz}\right)$ 和 $\det\left(\dfrac{dz}{dx}\right)$ 处处不为零.

如果区域 Ω_x 与 Ω_z 相同, 即 $\Omega_x = \Omega_z = \Omega$, 则称给出了区域 Ω 的一个变换. 因此, 区域 Ω 的变换就是在此区域中引进了新坐标, 使得此新坐标在此区域内可以处处通过旧坐标来表达, 反之亦然:

$$x^i = x^i(z^1, \cdots, z^n); \quad z^j = z^j(x^1, \cdots, x^n).$$

回想一下, 称一个元素的集合 G 为群是说, 如果在此集合中有两个运算: 对 G 中每对元素 g, h 对应有 G 中它们的积 $g \circ h$; 对 G 中每个元素 g 对应有 G 中的逆 g^{-1}. 它们应具有下列性质:

1) $(f \circ g) \circ h = f \circ (g \circ h)$;

2) 存在元 $1 \in G$, 使得 $1 \circ g = g \circ 1 = g$(群 G 的单位元);

3) $g \circ (g^{-1}) = 1$.

所给区域 Ω 上的所有变换构成一个群. 两个变换的积定义为: 如果 φ 为变换

$$x = x(z), \tag{1}$$

ψ 为变换

$$z = z(y), \tag{2}$$

则 $\varphi \circ \psi$ 为这两个变换的复合

$$x = x(z(y)). \tag{3}$$

对于逆变换 φ^{-1} 我们定义为:

$$z = z(x) \tag{4}$$

(即我们从等式 (1) 把坐标 z^j 反过来通过 $x^1,\cdots,x^n$ 表出). 恒同变换

$$x^i = z^i, \quad i = 1,\cdots,n \tag{5}$$

起了单位元 1 的作用. 不难验证性质 1),2),3).

从全体变换形成的群 G 中我们可以进一步选出一些对几何重要的子群. 假设区域 Ω 中有某个度量 (黎曼的或伪黎曼的), 它在坐标 $x^1,\cdots,x^n$ 下由对称的非异矩阵 $g_{ij} = g_{ji}(x^1,\cdots,x^n)$ 表示. 如果给定变换 $x^i = x^i(z^1,\cdots,z^n)$, 那么在坐标 $z^1,\cdots,z^n$ 下这个矩阵成为矩阵 $g'_{ij} = g'_{ij}(z^1,\cdots,z^n)$, 即

$$g'_{ij} = \frac{\partial x^k}{\partial z^i} g_{kl} \frac{\partial x^l}{\partial z^j}. \tag{6}$$

定义 4.1. 称变换 $x^i = x^i(z^1,\cdots,z^n)$ 是该度量下的一个运动 ① 是说, 如果有

$$g'_{ij}(z^1,\cdots,z^n) = g_{ij}(x^1(z),\cdots,x^n(z)). \tag{7}$$

因此, 度量运动就是保持内积 (g_{ij}) 的形式不变. 成立下面简单的命题.

引理 4.1. 所给度量下的所有运动构成一个群.

事实上, 如果变换 φ 和 ψ 保持度量形式不变, 则它们的复合也保持了度量形式, 逆变换 φ^{-1} 也如此. 此外, 由定义知恒同变换保持度量形式.

称这个群为该度量下的运动群.

2. 平面的变换

a) 设 x^1, x^2 是平面上的笛卡儿坐标. 平面变换群的最简单的例子是平移, 它是把整个平面沿某个向量 $\xi = (\xi^1,\xi^2)$ 移动. 在坐标表示下这个平移具形式

$$x^1 = z^1 + \xi^1, \quad x^2 = z^2 + \xi^2. \tag{8}$$

两个按向量 ξ 和 η 平移的积为

$$x^1 = z^1 + (\xi^1 + \eta^1), \quad x^2 = z^2 + (\xi^2 + \eta^2),$$

即仍旧是一个按向量 $\xi + \eta$ 的平移, 变换 (8) 的逆变换为

$$z^1 = x^1 - \xi^1, \quad z^2 = x^2 - \xi^2. \tag{9}$$

这也是按向量 $-\xi$ 的平移. 恒同变换是按零向量的平移. 因此平面中的平移构成了一个群. 每个平移对应了平面中一个向量 ξ, 平移的乘积对应于向量的和, 平移的逆对应 $-\xi$. 用群论的语言表示, 称保持群运算的一一对应为同构. 这表明所有平面的平移同构于平面向量的群. 这个群是交换的 (阿贝尔的), 即 $\xi + \eta = \eta + \xi$.

①我们常说的这个空间的运动意味着保持所给度量不变, 即保距变换.

b) 下面的例子是平面的伸缩(或相似扩大), 它一般说来并不是运动. 用坐标表示, 伸缩由公式

$$x^1 = \lambda z^1, \quad x^2 = \lambda z^2 \tag{10}$$

定义, 其中 λ 为任意的非零实数. 任意两个各以因子 λ 和 μ 倍拉伸的伸缩的积则具有形式

$$x^1 = \lambda\mu y^1, \quad x^2 = \lambda\mu y^2. \tag{11}$$

(10) 的逆伸缩为

$$z^1 = \frac{x^1}{\lambda}, \quad z^2 = \frac{x^2}{\lambda}. \tag{12}$$

因此平面的伸缩构成一个群. 这个群也是阿贝尔的, 并同构于非零实数在乘积下的群.

c) 平移连同伸缩: 这里指的是由公式

$$x^1 = \lambda z^1 + \xi^1, \quad x^2 = \lambda z^2 + \xi^2, \quad \lambda \neq 0. \tag{13}$$

给出的变换. 如果 $z^1 = \mu y^1 + \eta^1, z^2 = \mu y^2 + \eta^2$ 是另一个这种变换, 则这两个变换的复合具有形式

$$\begin{aligned} x^1 &= (\lambda\mu)y^1 + (\xi^1 + \lambda\eta^1), \\ x^2 &= (\lambda\mu)y^2 + (\xi^2 + \lambda\eta^2). \end{aligned} \tag{14}$$

因此, 第一个变换由偶对 (λ, ξ) 定义, 其中 λ 为非零实数, ξ 为平面中的向量, 第二个变换由偶对 (μ, η) 定义, 而它们的复合 $(\lambda, \xi) \circ (\mu, \eta)$ 为偶对 $(\lambda\mu, \xi + \lambda\eta)$. 如果我们将 (13) 等同于 (λ, ξ) (其中 λ 为非零实数, 而 ξ 为平面的向量), 于是复合 $(\lambda, \xi) \circ (\mu, \eta)$ 由公式

$$(\lambda, \xi) \circ (\mu, \eta) = (\lambda\mu, \xi + \lambda\eta) \tag{15}$$

定义. (λ, ξ) 的逆变换可写为 $\left(\frac{1}{\lambda}, -\frac{1}{\lambda}\xi\right)$ 的形式. 因此平移和伸缩一起构成了一个群. 这个群是非阿贝尔的. 事实上, 由公式 (15) 得出复合 $(\mu, \eta) \circ (\lambda, \xi)$ 等于

$$(\mu, \eta) \circ (\lambda, \xi) = (\lambda\mu, \eta + \mu\xi) \neq (\lambda, \xi) \circ (\mu, \eta).$$

注 平面的平移和伸缩构成的群有一个正规子群 (平移群) 和一个同构于伸缩群的商群. 公式 (15) 定义了平面平移群和非零实数乘法群的半直积, 后者是由伸缩构成的.

d) 平面的线性变换. 这个变换有形式

$$\begin{aligned} x^1 &= az^1 + bz^2, \\ x^2 &= cz^1 + dz^2 \end{aligned} \quad \text{或者} \quad \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} z^1 \\ z^2 \end{pmatrix}. \tag{16}$$

这种变换由矩阵 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 定义. 为了使变换 (16) 可逆, 即要从等式 (16) 中能用 x^1, x^2 表达出 z^1, z^2, 需要行列式 $\Delta = ad - bc$ 不为零, 即矩阵 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 非异. 如果有另一个非异矩阵 $\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$ 和

$$\begin{aligned} z^1 &= a'y^1 + b'y^2, \\ z^2 &= c'y^1 + d'y^2 \end{aligned} \tag{17}$$

对应的线性变换, 那么 (16) 和 (17) 的复合有形式

$$\begin{aligned} x^1 &= (aa' + bc')y^1 + (ab' + bd')y^2, \\ x^2 &= (ca' + dc')y^1 + (cb' + dd')y^2. \end{aligned} \tag{18}$$

这个变换也是线性的, 它对应于矩阵 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 和 $\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$ 的乘积. 因此, 平面的线性变换群同构于二阶非异矩阵群. 这个群也是非阿贝尔的.

e) 仿射群, 它由线性变换群与平移组合而成. 这个群的每个变换由偶对 (A, ξ) 决定, 其中 A 为非异矩阵, ξ 为平面向量. 仿射变换的形式为

$$\begin{aligned} x^1 &= az^1 + bz^2 + \xi^1, \\ x^2 &= cz^1 + dz^2 + \xi^2 \end{aligned} \quad \text{或者} \quad x = Az + \xi, \tag{19}$$

其中

$$\Delta = ad - bc \neq 0.$$

仿射群中的乘法规则是

$$(A, \xi) \circ (B, \eta) = (AB, \xi + A\eta). \tag{20}$$

仿射群是二阶非异矩阵群和平面向量群的半直积. 可以证明任何一个欧氏空间的运动都是仿射变换 (参看节尾的习题 4.2).

f) 假设在平面给出了欧氏度量, x^1, x^2 为平面上的欧氏坐标. 在此欧氏坐标下矩阵 g_{ij} 具形式

$$g_{ij} = \delta_{ij}.$$

我们来解释哪些仿射变换 (19) 是欧氏度量下的运动. 在坐标 (z^1, z^2) 下度量的矩阵是 g'_{ij}, 其中

$$g'_{ij} = \frac{\partial x^k}{\partial z^i}\delta_{kl}\frac{\partial x^l}{\partial z^j} = \sum_{k=1}^{2}\frac{\partial x^k}{\partial z^i}\frac{\partial x^k}{\partial z^j}. \tag{21}$$

仿射变换的雅可比矩阵 $\left(\frac{\partial x^k}{\partial z^i}\right)$ 与矩阵 $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 相同. 如果仿射变换 (19) 为运动, 则 $g'_{ij}=\delta_{ij}$, 这个等式转换为三个方程的方程组

$$a^2+c^2=1,\quad ab+cd=0,\quad b^2+d^2=1. \tag{22}$$

或者写成矩阵形式为

$$A^{\mathrm{T}}A=1.$$

这表明矩阵 A 为正交.

因此,(19) 为欧氏度量下的运动当且仅当矩阵 A 为正交矩阵. 方程 (22) 可以被显式地解出. 因为 $a^2+c^2=1$, 故可以用角 $\varphi, 0\leqslant\varphi<2\pi$ 表出, 使 $a=\cos\ \varphi, c=\sin\ \varphi$. 因而

$$d=\cos\ \varphi, b=-\sin\ \varphi \quad 或 \quad d=-\cos\ \varphi,\ \ b=\sin\ \varphi.$$

对每个 φ 我们得到两个正交矩阵类型

$$A=\begin{pmatrix} \cos\ \varphi & \sin\ \varphi \\ -\sin\ \varphi & \cos\ \varphi \end{pmatrix}, \tag{23}$$

$$A=\begin{pmatrix} \cos\ \varphi & \sin\ \varphi \\ -\sin\ \varphi & \cos\ \varphi \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{24}$$

第一类矩阵给出了整个平面以角 φ 绕坐标原点的旋转. 这类矩阵的行列式等于 1. 第二类变换化为两个变换的复合: 以角 φ 进行旋转, 再对横轴反射

$$z^1=y^1,\quad z^2=-y^2.$$

这时矩阵 A 的行列式等于 -1. 第一种类型是个矩阵行列式为 $+1$ 的运动, 它构成了整个运动群的一个子群. 这样的运动 (即无反射的运动) 被称做*正常运动*.

引理 4.2. a) 任意一个平面的正常运动是绕某点的旋转或者是平移.

b) 任一个形如 $z\mapsto Az+\xi$ 的运动, 其中矩阵 A 的行列式为 -1, 是对某条直线的反射和沿着此反射轴平移的复合 (或称滑动反射).

证明 设运动为

$$x=Az+\xi,\quad A=\begin{pmatrix} \cos\ \varphi & \sin\ \varphi \\ -\sin\ \varphi & \cos\ \varphi \end{pmatrix}, \tag{25}$$

即 $z\mapsto Az+\xi$. 如果 $\varphi=0$, 则 $A=1$, 从而得到平移. 设 $A\neq 1$. 我们来找出这个运动的旋转中心. 这个中心应是变换 (25) 的不动点, 即点 z_0 使

$$z_0=Az_0+\xi. \tag{26}$$

由 (26) 得到

$$(1-A)z_0=\xi. \tag{27}$$

如果形如 (25) 的矩阵 A 不为 1, 则矩阵 $1-A$ 必定非退化, 因此由等式 (27) 唯一地确定了旋转中心 z_0. 断言 a) 得证.

现在考虑 $\det A=-1$. 的平面运动群. 当绕轴的旋转使得 (参看公式 (24))

$$A=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \tag{28}$$

即 A 为对横轴反射的矩阵. 变换 $z\mapsto Az+\xi$ 有形式

$$\begin{pmatrix}z^1\\ z^2\end{pmatrix}\longmapsto\begin{pmatrix}z^1+\xi^1\\ -z^2+\xi^2\end{pmatrix}. \tag{29}$$

令

$$z^1=y^1,\quad z^2=y^2+\frac{1}{2}\xi^2 \tag{30}$$

使原来的坐标改变. 在此平移后我们的运动具形式

$$\begin{pmatrix}y^1\\ y^2\end{pmatrix}\longmapsto\begin{pmatrix}y^1+\xi^1\\ -y^2\end{pmatrix}. \tag{31}$$

这是一个滑动反射. 在特殊情形 $\xi^1=0$ 时, 我们得到纯粹的反射. 引理得证. □

因此, 我们得到三种类型的平面运动: 平移, 旋转和滑动反射 (特殊情形为反射).

g) 在这里我们考虑运动和伸缩生成的群. 我们要弄清楚在变换

$$x=\lambda Bz+\xi,\quad 其中 B^{\mathrm{T}}B=1 \tag{32}$$

下欧氏度量发生了什么情况. 这时的雅可比矩阵 $A=\left(\dfrac{\partial x}{\partial z}\right)$ 的形状为 $A=\lambda B$, 其中 B 为正交矩阵. 于是在坐标 (z^1,z^2) 下此欧氏度量的形式为

$$g'_{ij}=\lambda^2 g_{ij}=\lambda^2\delta_{ij}. \tag{33}$$

因此, 在这些变换下矩阵 g_{ij} 都被乘以同一个数. 这样的仿射变换被称做共形变换. 我们要弄清仿射变换 (19) 中哪些是共形的. 像前面那样进行论证, 我们得到矩阵 A 是共形的条件

$$\left(\frac{1}{\lambda}A^{\mathrm{T}}\right)\left(\frac{1}{\lambda}A\right)=1. \tag{34}$$

于是, 矩阵 $B=\dfrac{1}{\lambda}A$ 为正交. 这样, 共形子群由形如

$$x=\lambda Bz+\xi \tag{35}$$

的仿射变换组成, 其中 B 为正交矩阵

假设矩阵 B 的行列式等于 $+1$, 即 B 具有形式

$$B=\begin{pmatrix}\cos\ \varphi & \sin\ \varphi\\ -\sin\ \varphi & \cos\ \varphi\end{pmatrix}.$$

引进复变量 $v=z^1+iz^2, w=x^1+ix^2$. 于是由等式 (35) 我们得到

$$w=(x^1+ix^2)=\lambda(\cos\ \varphi-i\sin\ \varphi)(z^1+iz^2)+(\xi^1+i\xi^2),$$

即

$$w=\alpha v+\beta, \tag{36}$$

这里的复数 α 和 β 具形式

$$\alpha=\lambda e^{-i\varphi},\quad \beta=\xi^1+i\xi^2. \tag{37}$$

公式 (37) 表明正常 (即 $\det\ B=1$) 的平面共形变换就是复平面 $\mathbb{C}=\mathbb{C}^1$ 的复线性变换

$$v=z^1+iz^2\mapsto w=x^1+ix^2.$$

假设矩阵 B 的行列式等于 -1. 我们看出 (参见上面的例题) 这样的变换是由旋转再加上对横轴的反射得到, 这个反射为

$$x^1=z^1,$$

$$x^2=-z^2.$$

用复变量 v 和 w, 此变换可以表达为

$$w=\bar{v}=z^1-iz^2. \tag{38}$$

因此, 平面上欧氏度量的共形仿射变换的一般形式为

$$w=\alpha v+\beta\quad 或\quad w=\alpha\bar{v}+\beta, \tag{39}$$

其中 α 和 β 为任意复数且 $\alpha\neq 0$. 如果 $|\alpha|=1$, 变换 (39) 则为欧氏度量下的运动. 如果 α 为实数, 我们则得到了在例 c) 中分析过的平移加伸缩的变换.

3. 三维欧氏空间的运动

像在二维情形一样, 三维欧氏空间的运动都是仿射变换. 我们首先考虑作为线性变换并使坐标原点保持不动的运动. 它给出了三阶矩阵 $A=(a^i_j)$:

$$x=Az,\quad x^i=a^i_jz^j. \tag{40}$$

在坐标 x^1, x^2, x^3 下, 度量是欧氏的, 即 $g_{ij} = \delta_{ij}$. 这里的雅可比矩阵 $\left(\dfrac{\partial x^i}{\partial z^j}\right)$ 为矩阵 A. 于是在坐标 z^1, z^2, z^3 下度量为 g'_{ij}, 其中

$$g'_{ij} = a_i^k \delta_{kl} a_j^l = \sum_{i=1}^{3} a_i^k a_j^k.$$

如果 $g'_{ij} = \delta_{ij}$, 即变换 (40) 是个运动, 则有

$$\sum_{k=1}^{3} a_i^k a_j^k = \delta_{ij}. \tag{41}$$

等式 (41) 表示: 如果基向量 $e_1 = (1,0,0), e_2 = (0,1,0), e_3 = (0,0,1)$ 相互正交, 即 $\langle e_i, e_j \rangle = \delta_{ij}$, 则向量 $Ae_i = a_i^k e_i$ 也正交.

等式 (41) 的矩阵写法为

$$A^{\mathrm{T}} A = 1. \tag{42}$$

故而, 矩阵 A 为正交. 因为 $\det A^{\mathrm{T}} = \det A$, 故矩阵 A 的行列式等于 ± 1:

$$\det A = \pm 1. \tag{43}$$

正交矩阵的群与保持形式 $(x^1)^2 + (x^2)^2 + (x^3)^2$ 不变的变换群相同, 即是这样的矩阵 A 使得

$$\langle x, x \rangle = \langle Ax, Ax \rangle. \tag{44}$$

在这些条件下有

引理 4.3. 变换 A 有不变的直线, 它在 A 下或者完全不动或者是反射变换.

证明 设 v 为此直线的定向向量, 于是我们有

$$Av = \lambda v, \tag{45}$$

其中 λ 为实数. 于是 v 是具特征值 λ 的特征向量. 这里的 λ 是矩阵 A 的特征多项式的根:

$$\det(A - \lambda \cdot 1) = 0. \tag{46}$$

方程 (46) 关于 λ 是三次的, 并具有实系数. 因此它至少有一个实根 λ_0. 根 λ_0 对应于至少一个特征向量 v_0. 于是

$$\langle v_0, v_0 \rangle = \langle Av_0, Av_0 \rangle = \lambda_0^2 \langle v_0, v_0 \rangle.$$

由此知 $\lambda_0 = \pm 1$. 引理得证. □

设向量 w 垂直于特征向量 $v_0 : \langle w, v_0 \rangle = 0$. 于是向量 Aw 也垂直于 v_0:

$$\langle v_0, Aw \rangle = \pm \langle Av_0, Aw \rangle = \pm \langle v_0, w \rangle = 0. \tag{47}$$

结论 垂直于向量 v_0 并通过坐标原点的平面在变换 A 下不变.

在空间中选取欧氏坐标 x^1, x^2, x^3 使得 x^3 顺着向量 v_0 方向. 于是在此坐标下 A 有形式

$$A = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & \pm 1 \end{pmatrix}. \tag{48}$$

由等式 (42) 得出矩阵 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 为正交. 因而它具有形式

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \quad 或 \quad \begin{pmatrix} \cos\varphi & -\sin\varphi \\ -\sin\varphi & -\cos\varphi \end{pmatrix}. \tag{49}$$

于是按照 λ_0 和 $ad - cb$ 的符号我们有下列的运动类型.

a) *绕某条轴的旋转*. 以 x^3 轴为旋转轴, 这样的运动的矩阵有形式

$$A = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \det A = \lambda_0(ad - bc) = 1. \tag{50}$$

在 $\lambda_0 = -1, ad - bc = -1$ 的情形, 我们可以选取坐标 x^1, x^2, 使得矩阵 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 取形式 $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. 这时 $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ 为绕 x^1 轴旋转了角 π.

b) *反射旋转*. 这个变换由两个相继的变换产生: 绕某轴的旋转和对平面的反射, 其中的平面垂直于这条轴. 这时矩阵化为

$$A = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \det A = -1. \tag{51}$$

我们看到, 如果 $\det A = +1$ 则对应的正交变换总是绕某条轴的旋转. 正交矩阵 A 生成一个群, 以 $O(3)$ 表示. 具行列式 $+1$ 的正交矩阵生成 $O(3)$ 的子群, 记其为 $SO(3)$, $SO(3)$ 中每个矩阵至少有一个特征值 1. 由于已进行的讨论, $SO(3)$ 由经过坐标原点的所有可能的直线为轴的旋转组成.

现在我们来分析三维空间的任意运动. 如已经注意到的那样, 这些运动是仿射变换

$$z \mapsto Az + \xi. \tag{52}$$

完全像平面的情形一样, 我们确定矩阵 A 是正交的. 作到坐标原点的平移. 设 $z = y + y_0$. 于是在平移后, 这个运动可写为

$$y \mapsto Ay + (A-1)y_0 + \xi. \tag{53}$$

假设 A 具形式 (51), 即给出了反射旋转. 矩阵 $A-1$ 于是非异. 可以找到向量 y_0, 使得

$$(1-A)y_0 = \xi, \quad y_0 = (1-A)^{-1}\xi. \tag{54}$$

这时在坐标 y 下变换就是反射旋转:

$$y \mapsto Ay, \quad A = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{55}$$

现假设矩阵 $A \neq 1$ 具有 (50) 的形式, 它给出对某条轴的旋转. 这时矩阵 $A-1$ 退化. 关于 y_0 的方程 (54) 写成

$$(1-A)y_0 = \xi \leftrightarrow \begin{cases} (1-\cos\varphi)y_0^1 - \sin\varphi y_0^2 = \xi^1, \\ \sin\varphi y_0^1 + (1-\cos\varphi)y_0^2 = \xi^2, \\ \qquad\qquad\qquad\qquad\quad 0 = \xi^3. \end{cases} \tag{56}$$

如果 $\xi^3 \neq 0$, 此方程组无解. 但从前两个方程可以唯一地解出 y_0^1 和 y_0^2 (如果 $\varphi \neq 0$, 即变换 (52) 不是纯粹的平移). 这时在第三个坐标 y_0^3 的任意选取下, 运动 (52) 将具形式

$$y \mapsto Ay + (0,0,\xi^3), \quad A = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{57}$$

其中 $\eta = (0,0,\xi^3)$ 指向 A 的旋转轴, $A\eta = \eta$. 因此, 这个运动是螺旋的.

结论 三维空间的正常运动 ($\det A = +1$) 是螺旋运动(特别地, 是平移和旋转.)

螺旋旋转加上滑动反射得出整个运动群.

4. 变换群的其他例子

a) 按第 3 小节的分析方法, 我们考察 n 维欧氏空间的保持坐标原点不动的运动群. 以 $O(n)$ 记此群.$O(n)$ 中每个元由 n 阶正交矩阵 A 给出:

$$x = Az, \quad A^{\mathrm{T}}A = 1, \quad \det A = \pm 1. \tag{58}$$

形如 (58) 且 $\det A = 1$ 的变换构成子群 $SO(n) \subset O(n)$. 这个群也被称做 n 维旋转群. 然而这样的命名是有条件的, 因为当 $n \geqslant 3$ 时的 n 维旋转一般说来并不是平面的旋转.

群 $SO(n)$ 是连通的. 其意义如下: 如果 A_0 和 A_1 为任意两个旋转, $A_0, A_1 \in SO(n)$, 则在 $SO(n)$ 中存在一条曲线 $A(t), 0 \leqslant t \leqslant 1$ (即具行列式为 1 的正交矩阵的连续族), 使得 $A(0) = A_0, A(1) = A_1$.

只需对 A_0 为单位矩阵 $1 \in SO(n)$ 证明此断言. 由线性代数知道, 任意正交矩阵 (在这里是 A_1) 可以化为分块矩阵

$$\begin{pmatrix} \square & & & & & 0 \\ & \ddots & & & & \\ & & \square & & & \\ & & & \pm 1 & & \\ & & & & \ddots & \\ 0 & & & & & \pm 1 \end{pmatrix}, \tag{59}$$

其中第 i 块具形式

$$\square = \begin{pmatrix} \cos\ \varphi_i & \sin\ \varphi_i \\ -\sin\ \varphi_i & \cos\ \varphi_i \end{pmatrix}, \quad 0 \leqslant \varphi_i < 2\pi, \tag{60}$$

(在前一小节中对 $n = 2, 3$ 的情形得到过这种约化形式).

考虑矩阵族 $A(t)$, 在每个块 (60) 中将角 φ_i 换为 $t\varphi_i$. 当 $t = 1$ 时我们有矩阵 A_1, 而当 $t = 0$ 时得到对角线为 ± 1 的矩阵. 因为 $\det A = 1$, 于是 -1 的个数为偶数. 重新安排坐标, 将此矩阵化为下面形状.

$$A(0) = \begin{pmatrix} \square & & & & & 0 \\ & \ddots & & & & \\ & & \square & & & \\ & & & 1 & & \\ & & & & \ddots & \\ 0 & & & & & 1 \end{pmatrix}, \tag{61}$$

其中每一块有形状:

$$\square = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{62}$$

这样的矩阵可以经曲线与单位矩阵联结: 将每块 (62) 替换为

$$\begin{pmatrix} \cos\ t & \sin\ t \\ -\sin\ t & \cos\ t \end{pmatrix}. \tag{63}$$

在 $t=\pi$ 我们得到初始的矩阵, $t=0$ 时为单位矩阵.

显然, 群 $O(n)$ 不连通: 如果矩阵 A_0 和 A_1 各有 $\det A_0=1, \det A_1=-1$, 则这些矩阵不可能以曲线相连接. 事实上, 如果存在这样的曲线, 则 $\det A(t)$ 是 t 的连续函数. 但 $\det A_0=\det A(0)=1$ 而 $\det A(1)=-1$.

b) *伽利略群.* 经典力学中著名的伽利略相对性原理的论断是: 如果一个固定的参照系 (x^1,x^2,x^3) 置换到其他任意的参照系 (x'^1,x'^2,x'^3), 而后者相对于第一个系统作匀速的直线运动, 则所有经典力学的规律均保持其自身形式不变. 这就是说, 经典力学的规律相对于伽利略变换

$$\begin{aligned} t' &= t, \\ x'^1 &= x^1 - vt, \\ x'^2 &= x^2, \\ x'^3 &= x^3 \end{aligned} \tag{64}$$

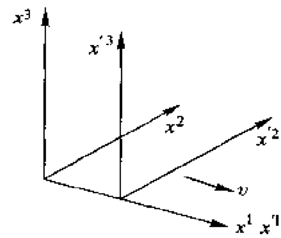


图 5 系统 (x^1,x^2,x^3) 固定不动, 系统 (x'^1,x'^2,x'^3) 以常速 v 沿第一条轴运动

不变. 这里的 v 为相对于固定系统的该系统的运动速度 (图 5). 称由从固定系统的变换 (64) 得到的参照系为*惯性参照系*. 我们作一个重要的评注: 在经典力学中, 时间 t (连同它的承载空间 $\mathbb{R}^3$) 具有绝对的属性, 就是说, 两个事件 A 和 B 之间的时间间隔量 Δt 不依赖于这个间隔是在哪个惯性参照系统中度量的. 换句话说, 时间的流动对整个宇宙是同一的; 在空间的每点上都放着同步的时钟. 从纯几何的观点看, 伽利略变换只不过是从一个坐标系 (t,x^1,x^2,x^3) 到另一个 (t',x'^1,x'^2,x'^3) 转移的公式, 其中 $t=t'$. 保持经典力学规律不变的伽利略群得出的一般变换形式为

$$\begin{aligned} t' &= t, \\ x' &= Ax + x_0 - vt, \end{aligned} \tag{65}$$

其中 A 为旋转矩阵.

定义 4.2. 称形如 (65) 的变换群为*伽利略群*.

力学中产生的不太显然的变换群的例子是

$$\begin{aligned} t &\to \beta t, \\ x &\to \alpha x; \end{aligned} \qquad \alpha^3 = \beta^2. \tag{66}$$

变换群 (66) 与开普勒第三定律相关, 这个定律说, 行星运转的周期的平方正比于与太阳距离的立方 (在近日点). 事实上, 开普勒第三定律是由这样的力学系统得出的, 即具势能 $\varphi=$ 常数 $/r$ 的牛顿引力场中的质点运动, 它关于变换 (66) 不变: 这些变换把运动轨道转换为运动轨道.

习题

4.1. 设 $Q(x) = b_{ij}x^i x^j$(其中 $b_{ij} = b_{ji}$) 为二次型.$B(x,y) = b_{ij}x^i y^j$ 为相应的双线性形式. 证明线性变换 A 保持双线性形式 $B(Ax, Ay) = B(x,y)$ 当且仅当它保持二次型不变:$Q(Ax) = Q(x)$ (向量 x 和 y 任意).

4.2. 证明欧氏空间的任意运动是仿射变换.

4.3. n 维欧氏空间中的仿射群同构于形如

$$\begin{pmatrix} A & \xi \\ 0 & 1 \end{pmatrix}$$

的 $n+1$ 阶矩阵的群, 其中 A 为非异 $n \times n$ 矩阵, ξ 为 n 维列向量.

4.4. 求伽利略群的矩阵表示.

§5. 弗莱纳公式

1. 平面曲线

考虑一个欧氏平面, 其欧几里得坐标为 x, y, 单位基向量为 e_1, e_2. 任意点 P 的坐标 (x,y) 表示了径向向量 $r = xe_1 + ye_2$, 它是由坐标原点 O 到 P 点的向量. 向量 r 的长由公式

$$|r| = \sqrt{\langle r, r\rangle} = \sqrt{x^2 + y^2} \tag{1}$$

给出. 假设有光滑曲线

$$r = r(t); \quad x = x(t), \quad y = y(t), \tag{2}$$

这里的曲线上点的径向向量为 $x(t)e_1 + y(t)e_2$. 此曲线的长为

$$l = \int_a^b \sqrt{(\dot{x})^2 + (\dot{y})^2}dt = \int_a^b dl, \tag{3}$$

其中 $\dot{x} = \dfrac{dx}{dt}, \dot{y} = \dfrac{dy}{dt}$; 这里的长度微分为 $dl = |v|dt, |v| = \sqrt{\langle v, v\rangle}, v = \dot{x}e_1 + \dot{y}e_2$ 是速度向量. 我们记 $v_t = \dfrac{dr}{dt}$, 显式地以 t 为下标, 表示是参照于参数 t 的速度向量. 用自然参数即长度 l 来考虑曲线常常是方便的 (参看 §2.1):

$$x = x(l), \quad y = y(l). \tag{4}$$

于是 $v = v_l = \dfrac{dx}{dl}e_1 + \dfrac{dy}{dl}e_2, |v_l| = 1$. 如果在曲线上给出了任意的参数 $t, x = x(t), y = y(t)$, 我们则有关系 $dl = \sqrt{\dot{x}^2 + \dot{y}^2}dt$. 有两个起重要作用的向量: 速度和加速度向量

$$\frac{dr}{dt} = v_t, \quad \frac{d^2r}{dt^2} = w_t. \tag{5}$$

如果是自然参数 $t=l$, 则 $|v_l|=1$. 我们有下面简单而常常有用的引理:

引理 5.1. 如果已知依赖于时间的向量 $v=v(t)$, 且 $|v|\equiv 1$, 则向量 v 和 $\dot{v}$ 垂直.

证明 因为 $v=v^1e_1+v^2e_2$ 且 $|v|^2=(v^1)^2+(v^2)^2$, 故

$$\begin{aligned}\frac{d}{dt}\langle v,v\rangle &= \langle \dot{v},v\rangle+\langle v,\dot{v}\rangle=2\langle v,\dot{v}\rangle\\ &=\frac{d}{dt}|v|^2=0;\end{aligned}$$

因此 $\langle v,\dot{v}\rangle=0$. 引理证完. □

注 如果有任意两个向量 $v(t)$ 和 $w(t)$, 则在欧氏几何中成立 $\frac{d}{dt}\langle v,w\rangle=\langle \dot{v},w\rangle+\langle v,\dot{w}\rangle$.

把我们的引理用于具自然参数 $l=t$ 的曲线 $r=r(t)=x(t)e_1+y(t)e_2$, 并设 $v=\frac{dr}{dt}$, 则得到

推论 在自然参数下速度 $v(t)$ 与加速度 $w(t)=\frac{dv}{dt}$ 相互垂直.

定义 5.1. 称加速度的值 $k=|w(t)|$ 为曲线 $r(t)$ 的曲率, 其中 $t=l$ 为自然参数. 称 $\hat{k}=\pm k$ 为定向曲率, 其中的符号与向量 v 和 w 的坐标构成的行列式的符号 (即标架 (v,w) 的定向的符号) 一致.

结论

$$\frac{dv}{dl}=kn=\frac{d^2r}{dl^2},\tag{6}$$

其中 n 为曲线的单位法向量:

$$n=\frac{w}{|w|}=\frac{1}{\sqrt{\left(\frac{d^2x}{dl^2}\right)^2+\left(\frac{d^2y}{dl^2}\right)^2}}\left(\frac{d^2x}{dl^2}e_1+\frac{d^2y}{dl^2}e_2\right).\tag{7}$$

称数 $1/k$ 为曲率半径R.

是否这个曲率概念与我们的直观观念一致? 曲率有如下性质:

1) 直线的曲率为零.

证明 设 $x=x_0+al, y=y_0+bl$ (直线). 取的是自然参数 l, 就是说

$$a^2+b^2=\left(\frac{dx}{dl}\right)^2+\left(\frac{dy}{dl}\right)^2=|v|^2=1.$$

因此 $w=\frac{d^2r}{dl^2}=0$, 从而 $k=0, R=\infty$. □

2) 半径为 R 的圆的曲率为常数 R^{-1}.

证明 设

$$x=x_0+R\cos\left(\frac{l}{R}\right),\quad y=y_0+R\sin\left(\frac{l}{R}\right),$$

其中 R 为常数. 于是

$$\frac{d^2x}{dl^2} = -\frac{\cos(l/R)}{R}, \quad \frac{d^2y}{dl^2} = \frac{\sin(l/R)}{R},$$

从而 $|w| = \dfrac{1}{R} = k$. □

有重要的定理:

定理 5.1. 如果以自然参数 l 表示的曲线 $r = r(l)$ 的曲率处处不为零, 则成立弗莱纳公式:

$$\begin{aligned} \frac{dv}{dl} &= w = kn, \\ \frac{dn}{dl} &= -kv, \end{aligned} \tag{8}$$

其中 $n = \dfrac{w}{|w|}$ 为单位法向量.

证明 由于 n 为单位法向量且 n 与 v 垂直, 我们有

a) $\left\langle n, \dfrac{dn}{dl} \right\rangle = 0$ (引理 5.1)

b) $\dfrac{dn}{dl} = \alpha v$ (因为 $n \perp v$, 且空间的维数为 2).

因为 $|v| = 1$, 于是 $|\alpha| = \left|\dfrac{dn}{dl}\right|$.$\alpha$ 等于什么? 由于 $\langle v, n\rangle = 0$, 我们有

$$\begin{aligned} 0 = \frac{d}{dl}\langle v, n\rangle &= \left\langle \frac{dv}{dl}, n \right\rangle + \left\langle v, \frac{dn}{dl} \right\rangle = k + \alpha\langle v, v\rangle \\ &= k + \alpha. \end{aligned}$$

因此 $\alpha = -k$. 定理得证. □

注 如果曲线的曲率在曲线的某点上为零, 那么可以在这点定义曲线的法向量 n, 使 $n \perp v$ 并使标架 (v, n) 为正定向. 这时把 k 换做定向曲率 $\tilde{k}$, 则弗莱纳公式 (8) 仍然成立 (请验证!).

弗莱纳公式的几何意义是什么? 由于 $\dfrac{dv}{dl} = kn, \dfrac{dn}{dl} = -kv$, 且 (v, n) 为单位法正交标架, 于是

$$v + \Delta v = v + (\Delta l)\frac{dv}{dl} + O(\Delta l^2) = v + k(\Delta l)n + O(\Delta l^2), \tag{9}$$

$$n + \Delta n = n + (\Delta l)\frac{dn}{dl} + O(\Delta l^2) = n - k(\Delta l)v + O(\Delta l^2).$$

令 $k(\Delta l) = \Delta\varphi$. 于是

$$\begin{aligned} \cos(\Delta\varphi) &= 1 + O(\Delta\varphi^2), \\ \sin(\Delta\varphi) &= \Delta\varphi + O(\Delta\varphi^2), \end{aligned} \tag{10}$$

从而我们可以写成

$$\begin{aligned} v + \Delta v &\approx \cos(\Delta\varphi)v + \sin(\Delta\varphi)n, \\ n + \Delta n &\approx -\sin(\Delta\varphi)v + \cos(\Delta\varphi)n, \end{aligned} \tag{11}$$

就是说, 由标架 (v, n) 到标架 $(v+\Delta v, n+\Delta n)$ 的改变是弯曲了一个小的角 $\Delta\varphi = k\Delta l$.

因此, 弗莱纳公式描述了在从点 l 到邻近点 $l+\Delta l$ 转移时标架 (v, n) 的弯曲, 它精确到一个二阶的小量. 常以公式

$$k = \left|\frac{d\varphi}{dl}\right|, \quad \hat{k} = \frac{d\varphi}{dl} \tag{12}$$

表示, 其中 φ 为向量 v(或 n) 在沿曲线运动时的旋转角.

现在提出一个自然的问题: 如果参数 t 不是自然的, 那么怎样计算平面曲线 $r(t) = (x(t), y(t))$ 的曲率?

此时 $v_t = (\dot{x}, \dot{y}), |v_t| \neq 1$. 因此向量 v_t 和 $\dot{v}_t = w_t$ (速度和加速度) 不垂直.

设 $\xi = \xi(t) = \xi^1 e_1 + \xi^2 e_2$ 为任意向量. 对我们的曲线而言有

$$dl = |v_t|dt = |\dot{r}|dt. \tag{13}$$

对于任意向量 $\xi(t)$ 我们有

$$\frac{d\xi}{dl} = \frac{1}{|v_t|}\frac{d\xi}{dt} = \frac{d\xi}{dt}\frac{dt}{dl}, \tag{14}$$

其中 $|v_t| = |\dot{r}|$, 而速度是按参数 t 沿曲线运动确定的.

设 $\eta(t) = \dfrac{v_t}{|v_t|} = \dfrac{dr}{dl}, \eta(t)$ 为单位切向量 (如果为自然参数, 则等于速度向量 v).

由曲率的定义, 我们有

$$k = \left|\frac{d^2r}{dl^2}\right| = \left|\frac{d\eta}{dl}\right| = \left|\frac{d}{dl}\left(\frac{v_t}{|v_t|}\right)\right| \tag{15}$$

(在自然参数下, 加速度的长等于曲率). 由公式 (14) 得出

$$\frac{d}{dl}\left(\frac{v_t}{|v_t|}\right) = \frac{1}{|v_t|}\frac{d}{dt}\left(\frac{v_t}{|v_t|}\right) = \frac{1}{|v_t|^2}\left(\frac{dv_t}{dt} - \frac{v_t}{|v_t|}\frac{d|v_t|}{dt}\right) = \frac{1}{|v_t|^2}\left(\ddot{r} - \frac{\dot{r}}{2|\dot{r}|^2}\frac{d|\dot{r}|^2}{dt}\right),$$

其中 $|v_t| = |\dot{r}|$. 故而我们得到 (假定 $|\dot{r}| \neq 0$)

$$w_l = \frac{d^2r}{dl^2} = \frac{d}{dl}\left(\frac{\dot{r}}{|\dot{r}|}\right) = \frac{1}{|\dot{r}|^2}\left(\ddot{r} - \frac{\langle\dot{r}, \ddot{r}\rangle}{|\dot{r}|^2}\dot{r}\right). \tag{16}$$

由此对于曲率有

$$k = \left|\frac{d^2r}{dl^2}\right| = \frac{1}{|\dot{r}|^2}|\eta - \langle\xi, \eta\rangle\xi|, \tag{17}$$

其中 $\eta = \ddot{r}, \xi = \frac{dr}{dl} = \frac{\dot{r}}{|\dot{r}|} = \frac{v_t}{|v_t|}$.

向量 $w_l = \frac{d^2r}{dl^2}$ 的分量有形式

$$w_l = \frac{1}{|\dot{r}|^2}\left(\ddot{x} - \dot{x}\frac{\dot{x}\ddot{x}+\dot{y}\ddot{y}}{\dot{x}^2+\dot{y}^2}\right)e_1 + \frac{1}{|\dot{r}|^2}\left(\ddot{y} - \dot{y}\frac{\dot{x}\ddot{x}+\dot{y}\ddot{y}}{\dot{x}^2+\dot{y}^2}\right)e_2. \tag{18}$$

进而

$$|w_l|^2 = k^2 = \frac{(\ddot{x}\dot{y}-\ddot{y}\dot{x})^2}{(\dot{x}^2+\dot{y}^2)^3}. \tag{19}$$

因此, 我们证明了关于曲率的重要公式.

定理 5.2. 对曲线 $x = x(t), y = y(t)$ 在任意选取的参数 t 下有公式

$$k = \left|\frac{d^2r}{dl^2}\right| = \frac{|\ddot{x}\dot{y}-\ddot{y}\dot{x}|}{(\dot{x}^2+\dot{y}^2)^{3/2}}, \tag{20}$$

其中假定 $\dot{x}^2+\dot{y}^2 = |\dot{r}|^2 \neq 0$.

注意, 分子是矩阵 $\begin{pmatrix}\dot{x} & \dot{y}\\ \ddot{x} & \ddot{y}\end{pmatrix}$ 行列式的绝对值.

2. 空间曲线. 曲率和挠率

用三维空间中的欧氏坐标给出的任意曲线 $x = x(t), y = y(t), z = z(t)$, 即 $r = r(t)$ 有

$$dl = |\dot{r}|dt = |v_t|dt = \sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}dt. \tag{21}$$

为了基本定义的简明起见, 我们就像在平面曲线的情形那样, 首先只考虑自然参数 $l, r = r(l), x = x(l), y = y(l), z = z(l)$. 由定义, $v = \dot{r} = \dot{x}e_1 + \dot{y}e_2 + \dot{z}e_3, w = \dot{v} = \ddot{r} = \ddot{x}e_1 + \ddot{y}e_2 + \ddot{z}e_3$ (这里的 "·" 表示对 l 的导数, 因为 $t = l$). 像在平面的情形, 我们引进

定义 5.2. 称相对于参数 l 的加速度的绝对值为空间曲线的曲率:$k = |w_l| = |\ddot{r}|$. 称曲率的倒数 $R = k^{-1}$ 为曲率半径.

我们由引理 5.1 知道, 速度和加速度向量相互正交:$v_l \perp w_l$, 这是因为 $|v_l| = 1$. 但是在三维情形, 向量 v_l 和 w_l 即便 $w_l \neq 0$ 也不够用来构成空间的一个完全的标架. 从直观上很显然, 在三维空间中仅一个曲率是不足以表示出曲线的几何性质的特性的. 例如, 绕在圆柱上的曲线 (圆螺旋线) 表示为

$$x = R\cos\ t, \quad y = R\sin\ t, \quad z = t.$$

除去曲率外它还向第三维方向扭转 (图 6). 第三个基向量可取为垂直于 v_l 和 w_l.

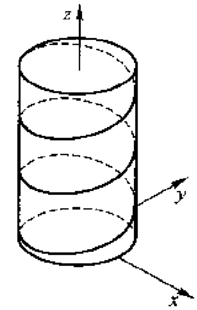


图 6

回忆在线性代数中，三维欧氏空间有重要的向量乘积的运算. 如果 ξ,η 为三维空间中的向量,

$$\xi=\xi^i e_i,\quad \eta=\eta^i e_i, \tag{22}$$

其中 e_i 为单位基向量 $(e_i\perp e_j,|e_i|=1)$, 则我们可以构造向量 $\gamma=[\xi,\eta],\gamma=\gamma^i e_i$, 其中 $\gamma^1=\xi^2\eta^3-\xi^3\eta^2,\gamma^2=\xi^3\eta^1-\xi^1\eta^3,\gamma^3=\xi^1\eta^2-\xi^2\eta^1$. 换句话说, $\pm\gamma^i$ 等于矩阵 $\begin{pmatrix}\xi^1 & \xi^2 & \xi^3\\ \eta^1 & \eta^2 & \eta^3\end{pmatrix}$ 的分部行列式, 即由去掉第 i 列后构成的矩阵行列式. 我们看出

$$\begin{aligned}&[\xi,\eta]=-[\eta,\xi],\quad [\xi_1+\xi_2,\eta]=[\xi_1,\eta]+[\xi_2,\eta],\\ &[\lambda\xi,\eta]=\lambda[\xi,\eta].\end{aligned} \tag{23}$$

还成立雅可比恒等式

$$[[\xi,\eta],\zeta]+[[\zeta,\xi],\eta]+[[\eta,\zeta],\xi]=0. \tag{24}$$

另外, 由解析几何知道向量积的性质:$[\xi,\eta]$ 指向向量 ξ,η 张成平面的垂直方向, 即与向量 $\lambda\xi+\mu\eta$ 垂直, 而且其长度等于

$$|[\xi,\eta]|=|\xi||\eta|\sin\ \varphi,$$

其中 φ 为 ξ 与 η 间的夹角,

$$\cos^2\ \varphi=1-\sin^2\ \varphi=\frac{\langle\xi,\eta\rangle^2}{|\xi|^2|\eta|^2}. \tag{25}$$

注 如果 ξ 和 η 在平面 (x,y) 上, 则它们的向量积正交于这个平面 (即指向 z 轴), 且 $[\xi,\eta]=(\xi^1\eta^2-\xi^2\eta^1)e_3$. 于是长度 $|[\xi,\eta]|$ 等于 $|\xi^1\eta^2-\eta^1\xi^2|=|\xi||\eta|\sin\ \varphi$.

在上一小节中所得到的表示平面曲线曲率的公式 (20) 现在可以重写为

$$k=\frac{|\ddot{x}\dot{y}-\ddot{y}\dot{x}|}{(\dot{x}^2+\dot{y}^2)^{3/2}}=\frac{|[\dot{r},\ddot{r}]|}{|\dot{r}|^3}, \tag{26}$$

它对任意参数 t 成立.

因此, 平面曲线曲率的一般公式可以通过向量积 $[\dot{r},\ddot{r}]$ 的长来表示这个曲率. 因为曲率与按弗莱纳公式的标架 (v,n) 旋转相关联, 于是标架 (v,n) 的绕正交于平面 (x,y) 的轴旋转的角速度向量自然地与曲率相关. 让我们回到空间曲线 $r=r(l),r=(x,y,z),x=x(l),y=y(l),z=z(l)$ 的情形. 对于它,

$$v(l)=\frac{dr}{dl},\quad w(l)=\frac{d^2r}{dl^2}. \tag{27}$$

我们假定 $|w| \neq 0$ 和 $|v| \neq 0$ (称这样的点为曲线的非异点). 假设 $|v| \equiv 1$; 于是 $\langle w, v\rangle = 0$, 即 $w \perp v$. 考虑向量 $b = [v, n]$, 其中 $n = \frac{w}{|w|}$. 称向量 n 为曲线的主法线, 向量 b 为次法线. 我们看到 $|b| = |v||n| \sin\varphi = 1, b \perp v, b \perp n$. 我们在 $w \neq 0$ 的曲线的点 (非异点) 上得到了法正交标架 (v, n, b).

如同内积情形那样, 我们要用到下面的论断:

引理 5.2. 设在三维空间中有两个向量 $\xi(t), \eta(t)$, 则成立莱布尼茨公式

$$\frac{d}{dt}[\xi, \eta] = \left[\frac{d\xi}{dt}, \eta\right] + \left[\xi, \frac{d\eta}{dt}\right]. \tag{28}$$

由对乘积函数微分的莱布尼茨公式, 即 $(fg)^{\cdot} = \dot{f}g + f\dot{g}$ 可立即得到它. 应该留意公式 (28) 右端中各乘积因子的次序.

定理 5.3. 对任意空间曲线 $r = r(l)$, 其中 l 为自然参数, 成立下面的弗莱纳公式

$$\begin{aligned} \frac{dv}{dl} &= kn, \\ \frac{dn}{dl} &= -\kappa b - kv, \\ \frac{db}{dl} &= \kappa n. \end{aligned} \tag{29}$$

这里的第一个等式直接定义了数 k, 而 $|\kappa| = \left|\frac{db}{dl}\right|$; 称数 κ 为空间曲线的挠率(它不必为正).

在平面情形 $b =$ 常向量, 从而 $|\kappa| = \left|\frac{db}{dl}\right| = 0$.

转向对空间曲线的弗莱纳公式的证明. 首先推导公式 $\frac{db}{dl} = \kappa n$. 因为 $b = [v, n]$, 故由引理 5.2 $[v, n]^{\cdot} = \dot{b} = [\dot{v}, n] + [v, \dot{n}]$. 由 $|v| = 1$ 和 $|n| = 1$, 则有 $\dot{v} \perp v, \dot{n} \perp n$. 故 $\dot{v} = kn, \dot{n} = \alpha v + \beta b$, 其中 α 和 β 为某些未知的常数. 因 $[n, n] = 0, [v, v] = 0$, 故 $[\dot{v}, n] = 0$, 且 $[v, \dot{n}] = \beta[v, b] = -\beta n$. 所以

$$\frac{db}{dl} = \frac{d}{dl}[v, n] = -\beta n = \kappa n, \tag{30}$$

其中 κ 由此等式决定.

因此, $\dot{b} = \kappa n$. 现计算 $\dot{n}$. 由于 $n = [b, v]$, 故

$$\frac{d}{dl}[b, v] = [\dot{b}, v] + [b, \dot{v}] = -\kappa b - kv.$$

定理得证. □

弗莱纳公式可以重写为在标架 (v, n, b) 下的矩阵形式. 记 $v = e_1, n = e_2, b = e_3$, 我们有

$$\frac{de_i}{dl} = b_i^j e_j \quad (i, j = 1, 2, 3), \tag{31}$$

矩阵 $B = (b_i^j)$ 具形式

$$B=\begin{pmatrix}0 & k & 0\\ -k & 0 & -\kappa\\ 0 & \kappa & 0\end{pmatrix}. \tag{32}$$

最后应该注意到曲率和挠率是欧氏空间中曲线的几何不变量的一个完全的集合. 更准确地说:

1) 如果已知平面曲线的函数关系 $k=k(l)$, 则可以准确到相差一个平面运动的范围内唯一地重新恢复此曲线. 称方程 $k=k(l)$ 为曲线的自然方程.

2) 如果已知空间曲线的函数 $k=k(l), \kappa=\kappa(l)$, 则可以准确到一个全空间运动的范围内重新恢复此曲线. 称这一对方程为空间曲线的自然方程.

对这两个定理的证明可参看 П.К. 拉舍夫斯基的书 [2].

3. 依赖于参数的正交变换

设 $A=(a_{ij}), i,j=1,\cdots,n$ 为正交矩阵 (参看 §4), 即

$$A^{\mathrm{T}}A=1 \Leftrightarrow \sum_{i=1}^{n} a_{ij}a_{ik}=\delta_{jk}, \tag{33}$$

其中每个 a_{ij} 都是参数 t 的函数. 又假设 $a_{ij}(0)=\delta_{ij}$, 即 $A(0)=1$. 我们有

引理 5.3. 如果 $A(0)=1$, 则矩阵 $B=\left.\left(\dfrac{dA}{dt}\right)\right|_{t=0}$ 为反称矩阵.

证明 对 (33) 式两端作关于 t 的微分得到

$$0=\frac{d}{dt}\delta_{jk}=\sum_{i=1}^{n}(\dot{a}_{ij}a_{ik}+a_{ij}\dot{a}_{ik}).$$

现令 $t=0$. 因为 $a_{ij}(0)=\delta_{ij}$, 故我们得到

$$0=\sum_{i=1}^{n}(\dot{a}_{ij}\delta_{ik}+\delta_{ij}\dot{a}_{ik})=b_{kj}+b_{jk},$$

这就是我们要证的. □

由此引理可以立即得到弗莱纳公式的新证明. 设 $e_1(l)=v(l), e_2(l)=n(l), e_3(l)=b(l)$ 为依赖于自然参数 l 的空间曲线上的标架. 于是从 $e_1(l), e_2(l), e_3(l)$ 经正交变换得到标架 $e_1(l+\Delta l), e_2(l+\Delta l), e_3(l+\Delta l)$, 这是因为这两个标架都是正交的

$$\begin{gathered} e_i(l+\Delta l)=\sum_{j=1}^{3}a_{ij}(l,\Delta l)e_j(l) \quad (i=1,2,3),\\ \sum_{i=1}^{3}a_{ij}(l,\Delta l)a_{ik}(l,\Delta l)=\delta_{jk}. \end{gathered} \tag{34}$$

对等式 (34) 作关于 Δl 的微分并令 $\Delta l = 0$, 我们得到

$$\dot{e}_i(l) = \sum_{j=1}^{3} b_{ij} e_j(l), \tag{35}$$

其中矩阵 (b_{ij}) 由引理 5.3 知为反称的. 因此矩阵 $B = (b_{ij})$ 有形式

$$B = \begin{pmatrix} 0 & b_{12} & b_{13} \\ -b_{12} & 0 & b_{23} \\ -b_{13} & -b_{23} & 0 \end{pmatrix}, \quad b_{ij} = -b_{ji}. \tag{36}$$

按定义

$$\dot{e}_1 = \frac{de_1}{dl} = \frac{dv}{dl} = kn = ke_2,$$

因此 $b_{12} = k, b_{13} = 0$. 这表明 $b_{21} = -k, b_{31} = 0$, 从而矩阵 B 具有如下形式:

$$B = \begin{pmatrix} 0 & k & 0 \\ -k & 0 & -b_{32} \\ 0 & b_{32} & 0 \end{pmatrix}.$$

记 $\kappa = b_{32}$, 我们便得到了弗莱纳公式 (29)(其矩阵形式 (31),(32)).

对于平面情形证明可类比地进行.

习题

5.1. 对于螺旋线 $r = (a\cos\ t, a\sin\ t, bt), a > 0, b \neq 0$, 求弗莱纳标架, 曲率和挠率.

5.2. 求下面曲线的曲率和挠率:

a) $r = e^t(\sin\ t, \cos\ t, 1)$;

b) $r = a(\mathrm{ch}\ t, \mathrm{sh}\ t, t)$.

5.3. 设椭圆的半轴等于 a 和 b, 求在其顶点处的曲率.

5.4. 求下列曲线的曲率和挠率:

a) $r = (t^2\sqrt{3}/2, 2 - t, t^3)$;

b) $r = (3t - t^3, 3t^2, 3t + t^3)$.

5.5. 证明: 如果曲线的曲率 $k(l) \equiv 0$, 则此曲线为直线.

5.6. 证明: 如果曲线的挠率 $\kappa(l) \equiv 0$, 则此曲线在平面中, 求此平面的方程.

5.7. 描述具有常曲率和常挠率的曲线类:$k(l) \equiv$ 常数 $\kappa(l) \equiv$ 常数.

5.8. 描述具常挠率的曲线类:$\kappa(l) =$ 常数.

5.9. 证明: 曲线 $r = r(t)$ 为平面的当且仅当混合积 $(\dot{r}, \ddot{r}, \dddot{r}) \equiv \langle \dot{r}, [\ddot{r}, \dddot{r}] \rangle = 0$.

5.10. 设 S 为平面曲线与一条 (平行于切线的) 和切线距离为 h 的割线之间的

面积. 用曲线的曲率表示出 $\lim\limits_{h\to 0}\dfrac{S^2}{h^3}$.

5.11. 证明: 对于光滑的闭曲线 $C: r=r(l)$ 有

$$\oint_C (r\dot{k} \quad \kappa b)dl = 0$$

5.12. 证明: 弗莱纳公式可以表示为

$$\dot{v}=[\zeta,v],\dot{n}=[\zeta,n],\dot{b}=[\zeta,b].$$

求向量 ζ(达布向量).

5.13. 解方程 $\dot{r}=[\omega,r],\omega=$ 常数.

5.14. 证明: 曲线 $r(l)$ 的曲率和挠率成比例: $k=c\cdot\kappa\ (k\neq 0),c$ 为常数当且仅当如果存在常向量 u 使得 $\langle u,v\rangle=$ 常数.

5.15. 假设一条曲线由向量 n 和 b 张成的法平面均通过一个固定点 x_0. 证明这样的曲线必在以此点为中心的球面上.

5.16. 证明: 曲线 $r(l)$ 在半径 R 的球面上当且仅当成立关系

$$R^2=\frac{1}{k^2}\left(1+\frac{\dot{k}^2}{(\kappa k)^2}\right).$$

5.17. 证明

$$\kappa=-\frac{(\dot{r},\ddot{r},\dddot{r})}{|[\dot{r},\ddot{r}]|}.$$

5.18. 对于平面曲线 $r=r(l)$, 我们考虑曲线 $n(l)$(n 为在所给点处的法线), l^* 为此曲线的自然参数. 证明

$$\frac{dl^*}{dl}=\sqrt{k^2+\kappa^2}.$$

5.19. 设

$$A=A(l)=\begin{pmatrix}0 & k(l) & 0\\ -k(l) & 0 & -\kappa(l)\\ 0 & \kappa(l) & 0\end{pmatrix}=(a_j^i(l)).$$

定义向量 $r_j=r_j(l)$ 为方程组

$$\frac{dr_j}{dl}=a_j^i r_i,\quad j=1,2,3,$$

其中 $r_1(0),r_2(0),r_3(0)$ 为给定的正交标架.

a) 证明: 对任意 l, 标架 $r_1(l),r_2(l),r_3(l)$ 为正交.

b) 假设 $r^*(l)=r_0+\displaystyle\int_0^l r_1(l)dl$. 证明: 这时 $r_1(l)=v^*(l),r_2(l)=n^*(l),r_3(l)=b^*(l)$, 其中 v^*,n^*,b^* 分别为曲线 $r^*(l)$ 的切线, 法线和次法线, 并且此曲线的曲率和挠率分别等于 $k(l),\kappa(l)$.

5.20. 设曲线在球面上并有常值曲率, 证明此曲线为圆.

§6. 伪欧几里得空间

1. 狭义相对论的最简单概念

回忆一下伪欧氏空间 $\mathbb{R}^n_{p,q}, p+q=n$, 它定义为一个有坐标 $x^1, x^2, \cdots, x^n$ 的空间, 其中向量 $\xi=(\xi^1,\cdots,\xi^n)$ 的"长度平方"由公式

$$|\xi|^2=\langle\xi,\xi\rangle=\sum_{i=1}^{p}(\xi^i)^2-\sum_{i=1}^{q}(\xi^{p+i})^2 \tag{1}$$

给出 (参看 §3.2). 在 $n=4, p=1$, 我们得到狭义相对论的时空空间 (闵可夫斯基空间 $\mathbb{R}^4_{1,3}=\mathbb{R}^4_1$), 其坐标为 x^0, x^1, x^2, x^3, 通常我们假定 $x^0=ct$, 其中常数 c 为真空中的光速. 我们也称空间 $\mathbb{R}^n_{1,n-1}=\mathbb{R}^n_1$ 为 (n 维) 闵可夫斯基空间.

在空间 $\mathbb{R}^4_1$ 中向量 $\xi=(\xi^0,\xi^1,\xi^2,\xi^3)$ 的长度平方由公式

$$|\xi|^2=\langle\xi,\xi\rangle=(\xi^0)^2-(\xi^1)^2-(\xi^2)^2-(\xi^3)^2 \tag{2}$$

给出. 数 $\langle\xi,\xi\rangle$ 可以为正, 为负也可为零. 使 $|\xi|=0$ 的向量 ξ 构成空间 $\mathbb{R}^4_1$ 中的锥, 我们称其为迷向锥面或光锥(参看图 7 的空间 $\mathbb{R}^3_1$ 的情形). 在锥内部的向量具有正的长度平方, $|\xi|^2>0$, 称其为类时向量. 在锥外部的向量具有负的长度平方, $|\xi|^2<0$, 称其为类空向量在图 7 中 ξ_+ 为类时向量, ξ_- 为类空向量, 而 ξ_0 在此迷向锥面上, 长度为零, 称这样的向量为迷向向量或光向量.

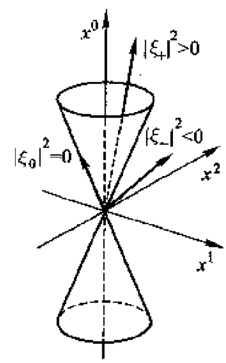


图 7 空间 $\mathbb{R}^3_1$ 中的迷向锥面:$(x^0)^2-(x^1)^2-(x^2)^2=0$

考虑任意质点的世界线 (参看 §1.1). 这条世界线在 $\mathbb{R}^4_1$ 中具形式

$$x^0=ct,\quad x^1=x^1(t),\quad x^2=x^2(t),\quad x^3=x^3(t). \tag{3}$$

这里的曲线 $x^1(t), x^2(t), x^3(t)$ 就是在三维空间 $\mathbb{R}^3$ 中通常的点的轨线. 切于世界线 (3) 的向量具有形式

$$\xi=(c,\dot{x}^1,\dot{x}^2,\dot{x}^3). \tag{4}$$

注意, $(\dot{x}^1,\dot{x}^2,\dot{x}^3)$ 是点在空间运动的速度向量 v 的坐标. 在狭义相对论中所用的假设是说质点的运动速度不能超过光速 c, 即 $|v|\leqslant c$. 这表示

$$c^2-(\dot{x}^1)^2-(\dot{x}^2)^2-(\dot{x}^3)^2\geqslant 0, \tag{5}$$

就是说向量 ξ 是类时的或者是迷向的. 特别, 如果我们的世界线就是光子的世界线, 那么向量 ξ 便是迷向的, 即 $|v|=c$. 由于这个理由, 我们也把迷向锥称为光

锥. 事实上, 只有无质量的点 (例如光子) 才能具有迷向的切向量. (正) 质点的世界线总具有类时的切向量. 特别地, (正) 质点的世界线整个严格地落在光锥的内部 (图 8, 注意, 在空间的所有点上都有迷向锥). 对于类时曲线 (即其切向量处处为类时的曲线) 可以像在欧几里得几何中那样定义长度的概念. 如果曲线具形式 $x^0 = x^0(\tau), x^1 = x^1(\tau), x^2 = x^2(\tau), x^3 = x^3(\tau), \xi = (\dot{x}^0, \dot{x}^1, \dot{x}^2, \dot{x}^3), |\xi|^2 > 0$, 则长度 l 具形式

$$l = \int_a^b |\xi| d\tau = \int_a^b \sqrt{(\dot{x}^0)^2 - \sum_{\alpha=1}^{3} (\dot{x}^\alpha)^2} d\tau. \tag{6}$$

在狭义相对论中, 量 l/c 被称做质点所耗去的固有时间. 像在欧几里得的情形那样, 我们有 $|v(l)| = 1$, 并称此参数 l 为类时世界线的自然参数.

如果点在三维空间中以匀速 $v = (v^1, v^2, v^3)$ 运动, 即

$$x^0 = ct, \quad x^1 = v^1 t, \quad x^2 = v^2 t, \quad x^3 = v^3 t, \tag{7}$$

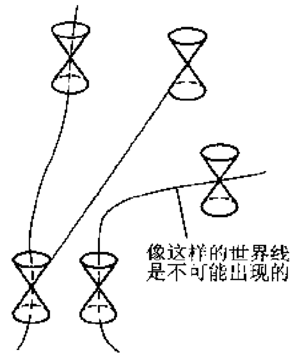


图 8 具质量或无质量的点的世界线

则有

$$dl = \sqrt{c^2 - v^2} dt = \sqrt{1 - \frac{v^2}{c^2}} dx^0,$$

$$l = x^0 \sqrt{1 - \frac{v^2}{c^2}}, \tag{8}$$

在这里我们采用了缩写的记号 $v^2 = |v|^2$.

特别, x^0/c 为质点 (在最初的坐标系中) 的固有时间.

2. 洛伦兹变换

我们在前面 (§4.4) 看到, 在经典 (牛顿) 力学中时间具有一种绝对的特性, 即在事件之间的时间区间 Δt 不依赖于对此区间进行测量的那个惯性参照系. 惯性参照系统通过伽利略变换 (4.64) 从一个系统得到另一个系统.

在狭义相对论中洛伦兹变换取代了伽利略变换. 转移到另一个参照系, 就是在空间 $\mathbb{R}^4_1$ 中选取新的坐标, 即闵可夫斯基空间的某个变换. 洛伦兹变换首先保持了坐标原点和时空区间不变, 即保持二次型 $dl^2 = c^2(dt)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2$. 到另一个惯性参照系的转移可以通过伪欧氏空间 $\mathbb{R}^4_1$ 的运动来实现. 称这些运动的整个群为*庞加莱群*.

先研究伪欧氏平面 $\mathbb{R}^2_1$ 的运动群. 首先假定这些运动保持坐标原点不动. 于是它由矩阵 A 给出

$$\begin{aligned} x^0 &= ax'^0 + bx'^1, \\ x^1 &= cx'^0 + dx'^1. \end{aligned} \tag{9}$$

如果坐标 x^0, x^1 是伪欧几里得的, 则其矩阵有形式 $g_{\alpha\beta} = \{g_{00} = 1, g_{11} = -1, g_{12} = g_{21} = 0\}$. 因为变换 (9) 为运动, 故

$$(g_{\alpha\beta}) = G = A^{\mathrm{T}} G A. \tag{10}$$

因为 $\det A^{\mathrm{T}} = \det A$, 且矩阵积的行列式等于行列式的积, 故有

$$(\det A)^2 = 1, \quad \det A = \pm 1.$$

对于矩阵 A 的分量, 由 (10) 得到由三个方程组成的方程组:

$$\mathrm{a}^2 - \mathrm{c}^2 = 1, \quad \mathrm{ab} - \mathrm{cd} = 0, \quad \mathrm{b}^2 - \mathrm{d}^2 = -1. \tag{11}$$

显然, $\mathrm{a} \neq 0$. 令 $\beta = \mathrm{c}/\mathrm{a}$. 直接的计算给出

$$\mathrm{a} = \pm\frac{1}{\sqrt{1-\beta^2}}, \quad \mathrm{c} = \pm\frac{\beta}{\sqrt{1-\beta^2}}, \quad \mathrm{b} = \pm\frac{\beta}{\sqrt{1-\beta^2}}, \quad \mathrm{d} = \pm\frac{1}{\sqrt{1-\beta^2}}. \tag{12}$$

这里 a 的符号与 c 的符号相同, 而 b 的符号同于 d 的符号. 于是, 作为空间 $\mathbb{R}^2_1$ 中伪欧氏度量的运动构成的所有变换 A 的群由下面形式的矩阵组成

$$A = \pm\begin{pmatrix} \dfrac{1}{\sqrt{1-\beta^2}} & \pm\dfrac{\beta}{\sqrt{1-\beta^2}} \\ \dfrac{\beta}{\sqrt{1-\beta^2}} & \pm\dfrac{1}{\sqrt{1-\beta^2}} \end{pmatrix}. \tag{13}$$

令 $\beta = \mathrm{th}\psi$, 引进双曲旋转角 ψ, 于是

$$A = \pm\begin{pmatrix} \mathrm{ch}\,\psi & \pm\mathrm{sh}\,\psi \\ \mathrm{sh}\,\psi & \pm\mathrm{ch}\,\psi \end{pmatrix}, \tag{14}$$

就是说, 我们得到了双曲旋转群. 任意双曲旋转把迷向锥 $|\xi|^2 = 0$ 转换到自身. 除此而外, 如果 $|\xi|^2 = 1$, 则 $|A\xi|^2 = 1$. 具有单位长的向量在 $\mathbb{R}^2_1$ 中构成了 "单位伪球面". 这个伪球面在 $\mathbb{R}^2_1$ 中由方程 $(x^0)^2 - (x^1)^2 = 1$ 给出, 从而就是一条双曲线 (图 9).

我们回想欧氏平面, 在那里的正交变换群由两个连通分支组成: 正常的和非正常的变换, 平面 $\mathbb{R}^2_1$ 的伪欧几里得运动群则更加复杂: 它们有四个连通分支 (由四片) 构成:

$$\begin{pmatrix} \mathrm{ch}\,\psi & \mathrm{sh}\,\psi \\ \mathrm{sh}\,\psi & \mathrm{ch}\,\psi \end{pmatrix}, \begin{pmatrix} \mathrm{ch}\,\psi & -\mathrm{sh}\,\psi \\ \mathrm{sh}\,\psi & -\mathrm{ch}\,\psi \end{pmatrix}, \begin{pmatrix} -\mathrm{ch}\,\psi & \mathrm{sh}\,\psi \\ -\mathrm{sh}\,\psi & \mathrm{ch}\,\psi \end{pmatrix}, \begin{pmatrix} -\mathrm{ch}\,\psi & -\mathrm{sh}\,\psi \\ -\mathrm{sh}\,\psi & -\mathrm{ch}\,\psi \end{pmatrix}. \tag{15}$$

第一个连通分支中的变换可以用曲线与单位 (即恒同) 变换相连接. 变换矩阵 $1, P, T, PT$ 各属于四个不同的连通分支, 它们的形状是:

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, P = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, T = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, PT = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

第一和第二个连通分支中的变换不改变时间 t 的方向 (这里的 $x^0 = ct$). 称这样的变换为*正时序变换*. 因此单位变换所在的连通分支由正常 (即行列式为 $+1$) 的正时序变换组成.

我们所考察的伪欧氏平面 $\mathbb{R}^2_1$ 的保持坐标原点不动的运动群被记为 $O(1,1)$. 这是空间 $\mathbb{R}^n_{p,q}$ 的伪正交变换的群 $O(p,q)$ 的特殊情形, 其中 $p+q=n$. 因此, $O(p,q)$ 为那些保持内积 (1) 不变的矩阵 A 的群

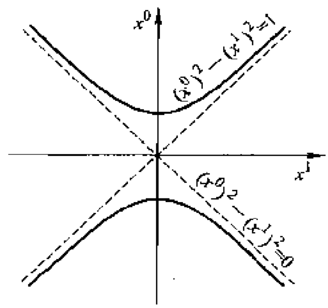


图 9

$$\langle A\xi, A\eta\rangle = \langle \xi,\eta\rangle = \xi^1\eta^1+\cdots+\xi^p\eta^p-\cdots-\xi^n\eta^n. \quad (16)$$

群 $O(1,n-1)$ 特别重要, 它是 n 维闵可夫斯基空间 $\mathbb{R}^n_1$ 的运动, 它保持原点不动. 我们已知群 $O(1,1)$ 由四片 (即四个连通分支) 组成. 事实上, 群 $O(1,n-1)$ 也由四片构成; 我们不在这里证明这个论断, 而只限于一个它的更弱的形式.

引理 6.1. *存在群 $O(1,n-1)$ 到群 $\mathbb{Z}_2\times\mathbb{Z}_2$ 的一个同态 φ* ①. *这里 $\mathbb{Z}_2\times\mathbb{Z}_2$ 是二阶整数群 $\mathbb{Z}_2=\{+1,-1\}$ 的直和.*

证明 设 e_0 为 x^0 轴的单位向量. 如果矩阵 $A\in O(1,n-1)$, 则令

$$\varphi(A) = (\det A, \operatorname{sgn}\langle e_0, Ae_0\rangle). \tag{17}$$

我们注意到内积 $\langle e_0, Ae_0\rangle$ 不为零, 这是因为向量 Ae_0 同时有

$$\langle Ae_0, Ae_0\rangle = \langle e_0, e_0\rangle = 1 \tag{18}$$

的缘故. 把映射 φ 的同态性留给读者去验证. □

称满足 $\varphi(A)=(1,1)$ 的变换 A 为*正常变换*. 对于使得 $\operatorname{sgn}\langle e_0, Ae_0\rangle=+1$ 的变换, 像前面那样, 称其为正时序的 (它们不改变时间 t 的方向). 同态 φ 的核② 由正常变换构成, 且是群 $O(1,n-1)$ 的单位元所在的连通分支. 我们不在此证明这些; 从所证明的引理得到的结论只是说, 此群中至少有四个连通分支.

伪欧氏空间 $\mathbb{R}^n_1$ 的整个运动群由群 $O(1,n-1)$ 再加上平移得到.

现在把已经得到的关于伪欧氏空间的运动的结果转换成狭义相对论的语言. 正如已经说过的那样, 从一个惯性参照系 (ct,x^1,x^2,x^3) 到另一个 (ct',x'^1,x'^2,x'^3) 的转移是由保持二次形式 $(ct)^2-(x^1)^2-(x^2)^2-(x^3)^2$ 不变的变换来实现的. 假设系统 (x') 相对于 (x) 沿轴 x^1 以速度 v 运动. 这表明 $x^2=x'^2, x^3=x'^3$, 从而坐标的变换

①回忆从群 G_1 到群 G_2 的同态 φ 的概念. 这是一个变换 $\varphi: G_1\to G_2$, 使得 $\varphi(1)=1, \varphi(g_1g_2)=\varphi(g_1)\varphi(g_2), \varphi(g^{-1})=\varphi(g)^{-1}$. 常称 φ 为群 G_1 在 G_2 中的一个 "表示".

②从 G_1 到 G_2 的同态的核是群 G_1 中那些元素 g 的集合, 它满足 $\varphi(g)=1\in G_2$.

具形式

$$\begin{aligned} x^0 &= (ct) = \mathrm{a}(ct') + \mathrm{b}x'^1, \\ x^1 &= \mathrm{c}(ct') + \mathrm{d}x'^1; \quad x^2 = x'^2, \quad x^3 = x'^3. \end{aligned} \tag{19}$$

矩阵 $A = \begin{pmatrix} \mathrm{a} & \mathrm{b} \\ \mathrm{c} & \mathrm{d} \end{pmatrix}$ 在 $O(1,1)$ 中.

如果速度 v 减为 0, 则变换 (19) 变成恒等式, 故矩阵 A 应在群 $O(1,1)$ 的单位元所在的连通分支, 从而有形式

$$A = \begin{pmatrix} \mathrm{ch}\ \psi & \mathrm{sh}\ \psi \\ \mathrm{sh}\ \psi & \mathrm{ch}\ \psi \end{pmatrix}. \tag{20}$$

于是

$$\begin{aligned} ct &= ct'\mathrm{ch}\ \psi + x'^1\mathrm{sh}\psi, \\ x^1 &= ct'\mathrm{sh}\ \psi + x'^1\mathrm{ch}\ \psi. \end{aligned} \tag{21}$$

考虑有上标符号的系统的原点 O' 在系统 $K(ct, x^1)$ 中的运动. 当 $x'^1 = 0$, 故公式 (21) 有形式

$$ct = ct'\mathrm{ch}\ \psi,$$

$$x^1 = ct'\mathrm{sh}\ \psi,$$

或相除有 $x^1/ct = \mathrm{th}\ \psi$. 但显然 x^1/t 是系统 K' 相对于系统 K 的速度 v. 因此

$$\mathrm{th}\ \psi = \frac{v}{c}. \tag{22}$$

由此

$$\mathrm{sh}\ \psi = \frac{v/c}{\sqrt{1 - v^2/c^2}}, \quad \mathrm{ch}\ \psi = \frac{1}{\sqrt{1 - v^2/c^2}}. \tag{23}$$

代入 (21), 求得了

$$\begin{aligned} t &= \left(t' + \left(\frac{v}{c^2}\right)x'^1\right)\frac{1}{\sqrt{1 - v^2/c^2}}, \\ x^1 &= (x'^1 + vt')\frac{1}{\sqrt{1 - v^2/c^2}} \end{aligned} \tag{24}$$

称这个变换为洛伦兹变换.

假定速度 v 比光速 c 小得多, 即 $v/c \ll 1$, 由 (24) 式知道, 如果 $v/c \to 0$, 则洛伦兹变换便变成了伽利略变换 (4.64): $t = t', x^1 = x'^1 + vt$. 换句话说, 在参照系间有小的相对速度时, 相对论的公式便转化为经典力学的公式. 但在大速度 (与光速相比较而言) 的情形, 这两个理论有了深刻的差异. 我们在下述流行的相对论的结论上解释这些原则性的差异; 这些相对论的结论是在速度向量方向上线性尺度的减缩和时钟的时滞 (涉及同时在一个参照系的事件在另一参照系中有不同的时间).

设在系统 K 中放置了一根长为 l 的刚性棒, 它平行于 x^1 轴; 设其端点在 K 中的坐标分别为 x_1^1, x_2^1, 即 $l = x_2^1 - x_1^1$. 现在我们求这个棒① 在系统 K' 中的长度. 为此我们求其端点在系统 K' 中的坐标 $x_1'^1, x_2'^1$; 在瞬时 t' 我们有

$$x_1^1 = \frac{x_1'^1 + vt'}{\sqrt{1 - v^2/c^2}}, \quad x_2^1 = \frac{x_2'^1 + vt'}{\sqrt{1 - v^2/c^2}}. \tag{25}$$

在系统 K' 中棒长 l' 等于 $l' = x_2'^1 - x_1'^1$. 从 x_2^1 减去 x_1^1 我们得到

$$l = \frac{l'}{\sqrt{1 - v^2/c^2}}. \tag{26}$$

因此, 棒子在那个被放置的参照系中它具有最长的长度. 而在以速度 v 运动的系统中它的长按比例 $\sqrt{1 - v^2/c^2}$ 缩短 (洛伦兹收缩).

转而研究时间坐标 t. 设在系统 K 中有两个同时发生的事件 A_1 和 A_2(即在 K 中 $t(A_1) = t(A_2)$); 同时假定事件 A_1 发生在具坐标 (x_1^1, x_1^2, x_1^3) 的点上, 而 A_2 发生在坐标为 (x_2^1, x_2^2, x_2^3) 的点上, 其中 $x_1^1 \neq x_2^1$. 对处在系统 K 中的观察者来看, 在运动的参照系 K' 中发生了什么? 由洛伦兹变换公式 (24) 我们有

$$t(A_1) = \frac{t'(A_1) + (v/c^2)x_1'^1}{\sqrt{1 - v^2/c^2}} = t(A_2) = \frac{t'(A_2) + (v/c^2)x_2'^1}{\sqrt{1 - v^2/c^2}},$$

即 $t'(A_1) - t'(A_2) = \dfrac{v}{c^2}\sqrt{1 - \dfrac{v^2}{c^2}}(x_2'^1 - x_1'^1) \neq 0$. 所以 $t'(A_1) \neq t'(A_2)$, 即是说, 对在 K 中两个同时发生的事件发现在 K' 中是不同时的. 注意, 如果 $t(A_1) > t(A_2)$ 且 $x_2'^1 - x_1'^1$ 不小, 我们则得到 $t'(A_1) < t'(A_2)$, 就是说, 乍看起来似乎是结果先于原因. 建议读者自己理清这个臆想的悖论.

假定在系统 K' 中放置了一只时钟. 让我们取两个发生在系统 K' 中同一地 x'^1, x'^2, x'^3 中的事件, 分别记为事件 A_1 和 A_2. 在系统 K' 中这两个事件之间的时间为时间区间 $\Delta t' = t_2' - t_1'$. 我们现在来求在系统 K 中这些事件之间经过的时间 Δt. 由 (24) 式我们有

$$t_1 = \frac{t_1' + (v/c^2)x_1'^1}{\sqrt{1 - v^2/c^2}}, \quad t_2 = \frac{t_2' + (v/c^2)x_2'^1}{\sqrt{1 - v^2/c^2}},$$

或者, 从一个算出另一个,

$$t_2 - t_1 = \Delta t = \frac{\Delta t'}{\sqrt{1 - v^2/c^2}}. \tag{27}$$

我们得到 $\Delta t > \Delta t'$. 换句话说, 运动的时钟走得比固定的时钟慢, 即从系统 K 中的观察者来看, 系统 K' 中的时钟较之于他的时钟有一个滞后.

①我们要强调指出, 棒长不是四维空间的不变量而是三维投影的长.

我们来分析关于平行速度相叠加的重要问题. 设在系统 K' 中点 P 以速度 w' 沿 x'^1 轴作关于此系统的运动. 点 P 关于系统 K 的速度是什么? 显然, $w' = \dfrac{dx'^1}{dt'}, w = \dfrac{dx^1}{dt}$,

$$dt = \frac{dt' + (v/c^2)dx'^1}{\sqrt{1 - v^2/c^2}}, \quad dx^1 = \frac{vdt' + dx'^1}{\sqrt{1 - v^2/c^2}}, \quad w = \frac{dx^1}{dt} = \frac{v + w'}{1 + vw'/c^2}.$$

如果 $v, w' \ll c$, 则 $w \approx w' + v$, 于是我们得到了经典力学中速度叠加的通常公式. 我们注意, 如果 $w' = c$, 则对任意速度 $v < c$, 点 P 在系统 K 中的速度为 $\dfrac{v + c}{1 + vc/c^2} = c$, 即对光速 c 再加上任意的速度 v 并不改变光速.

习题

6.1. 在空间 $\mathbb{R}^3_{1,2}$ 中令

$$\xi \times \eta = (\xi^1\eta^2 - \xi^2\eta^1, \xi^0\eta^2 - \xi^2\eta^0, \xi^1\eta^0 - \xi^0\eta^1)$$

来定义 "向量积", 其中 $\xi = (\xi^0, \xi^1, \xi^2), \eta = (\eta^0, \eta^1, \eta^2)$.

a) 验证对于基向量 e_0, e_1, e_2 (e_0 为类时的) 的两两向量积有

$$e_0 \times e_1 = -e_2, e_0 \times e_2 = e_1, e_1 \times e_2 = e_0.$$

b) 证明: 向量积 "×" 为双线性反称运算, 同时成立雅可比恒等式:

$$\xi_1 \times (\xi_2 \times \xi_3) + \xi_3 \times (\xi_1 \times \xi_2) + \xi_2 \times (\xi_3 \times \xi_1) = 0.$$

c) 证明: 如此定义的向量积在 $O(1,2)$ 群中的正常变换下不变.

6.2. 设 $r = r(l)$ 为 $\mathbb{R}^3_{1,2}$ 中的类时曲线, 同时有 $(\dot r(l))^2 = (\dot r^0)^2 - (\dot r^1)^2 - (\dot r^2)^2 \equiv 1$, 且 $\dot r^0 > 0$. 引进向量 v, n, b, 其定义为 $v = \dot r, \dot v = kn, b = n \times v$. 证明弗莱纳公式的伪欧氏的类比:

$$\dot v = kn, \quad \dot n = kv - \kappa b, \quad \dot b = \kappa n.$$

6.3. 推导出引理 5.3 的关于在 $\mathbb{R}^3_{1,2}$ 中洛伦兹变换的导数的一个类比引理, 其中的变换依赖于参数.

6.4. 在 $\mathbb{R}^3_{1,2}$ 中解方程 $\dot r = \omega \times r$, 其中 ω 为一常向量.

6.5. 证明: 在 $\mathbb{R}^{n+1}_{1,n}$ 中类时向量的正交补是一个类空的超平面. 类空向量的正交补是什么? 光向量的呢?

第二章　曲面论

§7. 空间曲面的几何

1. 曲面上的坐标

三维空间中的曲面可以说是在其上出现了最简单的内蕴几何的对象. 这是什么意思呢?

我们研究过曲线以及它们在平面和空间中的度量不变量. 但是这些不变量 (曲率和挠率) 是曲线的所处在空间中状态的不变量, 这是外在几何的概念. 在曲线上不存在任何内在的度量不变量. 这表示, 沿曲线可以选取自然参数, 即在两点之间的曲线段的长, 其度量方法与度量直线段的完全一样:

$$l = \int_{t_0}^{t_1} |v_t| dt, \quad v_t = \dot{r} = (\dot{x}, \dot{y}, \dot{z}).$$

对曲面则不一样: 在球面上没有办法给出坐标 (甚至在部分球面上) 使得在此坐标下的长度公式能像在欧氏平面上用笛卡儿坐标给出的长度公式那样.

以什么样的方式给出曲面呢? 在三维空间中有三种办法给出曲面:

1) 最简单的: 作为函数

$$z = f(x, y)$$

的图像定义它.

2) 较一般的: 以方程

$$F(x, y, z) = 0$$

定义.

3) 参数的 (类似于曲线的):$r = r(u,v)$, 或更详细地, 写为 $x = x(u,v), y = y(u,v), z = z(u,v)$, 其中 u, v 在 (u,v) 平面中某个区域内变动.

定义 7.1. 称由方程 $F(x,y,z) = 0$ 给出的曲面在点 $P = (x_0, y_0, z_0)$ 为非异是说, F 在点 P 的梯度不为零

$$\frac{\partial F}{\partial x}e_1 + \frac{\partial F}{\partial y}e_2 + \frac{\partial F}{\partial z}e_3 \neq 0, \quad x = x_0, y = y_0, z = z_0,$$

其中 $F(x_0, y_0, z_0) = 0$.

由隐函数定理, 即如果 $\left.\frac{\partial F}{\partial z}\right|_{x_0,y_0,z_0} \neq 0$, 则可在点 $P = (x_0, y_0, z_0)$ 附近由方程 $F(x,y,z) = 0$ 解出 z, 就是说, 找到了函数 $z = f(x,y)$, 满足 $f(x_0, y_0) = z_0$ 并且 $F(x,y,f(x,y)) = 0$ 在点 (x_0, y_0) 附近的平面区域 (x,y) 中成立. 微分等式 $F(x,y,z) = 0$ 得到

$$\frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy + \frac{\partial F}{\partial z}dz = 0,$$

并且由此表明

$$\frac{\partial f}{\partial x} = -\frac{\partial F/\partial x}{\partial F/\partial z}, \quad \frac{\partial f}{\partial y} = -\frac{\partial F/\partial y}{\partial F/\partial z}.$$

于是, 由方程 $F(x,y,z) = 0$ 给出的曲面在非异点的邻域内可以以图像的形式给出. 所以在非异点附近, 总可以用参数给出曲面: $z = f(u,v), x = u, y = v$(在点 $x_0 = u_0, y_0 = v_0$ 附近). 换句话说, 在非异点附近可以给出局部坐标 (u,v).

反之, 设曲面由参数给出

$$r = r(u,v); \quad x = x(u,v), \quad y = y(u,v), \quad z = z(u,v).$$

定义 7.2. 称点 $P = (x_0, y_0, z_0) = (x(u_0, v_0), y(u_0, v_0), z(u_0, v_0))$ 为非异是说矩阵

$$A = \left.\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \end{pmatrix}\right|_{u_0,v_0}$$

的秩为 2.

定理 7.1. 如果曲面由参数给出, 并且点 $P = (u_0, v_0)$ 为非异, 则在此点附近曲面可给出方程 $F(x,y,z) = 0$, 其中 $F(x_0, y_0, z_0) = 0$ 且 $\operatorname{grad} F|_{x_0,y_0,z_0} \neq 0$.

证明 由非异点的定义知矩阵 A 的秩为 2. 为确定起见, 设行列式

$$\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u} \neq 0.$$

回忆逆映射定理. 设 $x = x(u,v), y = y(u,v)$, 在点 (u_0, v_0) 的雅可比不为零, 并且 $x_0 = x(u_0, v_0), y_0 = y(u_0, v_0)$. 于是在点 (x_0, y_0) 的某个邻域内可以找到逆映射

$$u = u(x,y), \quad u_0 = u(x_0, y_0), \quad v = v(x,y), \quad v_0 = v(x_0, y_0),$$

同时矩阵 $\begin{pmatrix} u_x & v_x \\ u_y & v_y \end{pmatrix}$ 是矩阵 $\begin{pmatrix} x_u & y_u \\ x_v & y_v \end{pmatrix}$ 的逆. 利用这个定理我们求出表达式 $u(x,y), v(x,y)$; 然后, 把它们代入 z 的表达式中得到函数 $z(u(x,y),v(x,y))$, 并且 $z_0 = z(u_0,v_0) = z(u(x_0,y_0),v(x_0,y_0))$. 我们得到在非异点 x_0,y_0,z_0 附近曲面局部地表示为图像的形式 $z = f(x,y), z_0 = f(x_0,y_0)$. 定理得证. □

结论是, 局部地, 在曲面的非异点 $P = (x_0,y_0,z_0)$ 附近, 所有这三种给出曲面 (以光滑函数表示) 的方式相互等价.

例

7.1. $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$ (椭球面). a) 无奇点; b) 整体上不能由图像给出 (局部则可以);c) 整体上不能由参数给出 (为此, 所有点都是非异的).

7.2. $\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$ (单叶双曲面). a) 整体上不能由图像给出; b) 整体上可由参数给出, 选取的参数可以是 $u=z, v=\varphi$, 其中 φ 为极角.

7.3. $-\frac{x^2}{a^2}-\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$ (双叶双曲面). 每一叶可以由图像 $z=f(x,y)$ 给出, 因而可参数化, 并因而所有点为非异.

7.4. 锥面 $\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0$. 点 $(0,0,0)$ 为奇点.

现在考察一般的情形, 假设曲面由 n 维空间中的区域中的方程组给出:

$$f_1(x^1,\cdots,x^n)=0,\cdots,f_{n-k}(x^1,\cdots,x^n)=0. \tag{1}$$

定义 7.3. 称曲面 (1) 的点 $(x_0^1,\cdots,x_0^n)$ 为非异是说, 矩阵 $\left(\frac{\partial f_i}{\partial x^j}\right)_{x^j=x_0^j}$, $i=1,\cdots,n-k, j=1,\cdots,n$ 的秩等于 $n-k$.

有下面的简单论断.

引理 7.1. 在曲面(1) 的非异点$(x_0^1,\cdots,x_0^n)$ 附近可以引进局部坐标$(z^1,\cdots,z^k)$.

证明 假设行列式不等于零的 $n-k$ 阶子式为 $\left(\frac{\partial f_i}{\partial x^q}\right)_{x^q=x_0^q}$, $q=k+1,\cdots,n$. 于是在曲面的点 $(x_0^1,\cdots,x_0^n)$ 的邻域中我们有缺少了一些变量的 $(x^1,\cdots,x^k)$ 作为局部坐标. 在此点的邻域中解方程组 (1), 由隐函数定理得到

$$\begin{gathered} x^{k+1}=x^{k+1}(z^1,\cdots,z^k), \\ \cdots\cdots\cdots\cdots \\ x^n=x^n(z^1,\cdots,z^k), \\ z^1=x^1,\cdots,z^k=x^k. \end{gathered}$$

我们得到了在所讨论的点的邻域中曲面用变量 $(z^1,\cdots,z^k)$ 给出的参数表示. □

n 维空间中的超曲面是个重要的特殊情形. 它们由一个方程给出:

$$f(x^1, \cdots, x^n) = 0. \tag{2}$$

在超曲面 (2) 上的点 $(x_0^1, \cdots, x_0^n)$, 即使得 $f(x_0^1, \cdots, x_0^n) = 0$ 的点为非异表明

$$\text{grad } f = \left(\frac{\partial f}{\partial x^1}, \cdots, \frac{\partial f}{\partial x^n}\right)\bigg|_{x^j = x_0^j} \neq 0. \tag{3}$$

2. 切平面

设 $\mathbb{R}^3$ 中一个曲面由参数给出:$r = r(u, v), r = (x, y, z)$, 其中 u, v 为曲面上的坐标. 以坐标 u, v 给出的曲线 $u = u(t), v = v(t)$ 定义了我们在 $\mathbb{R}^3$ 中曲面上的曲线 $r(t) = r(u(t), v(t))$. 它的速度向量为

$$\dot{r}(t) = r_u\dot{u} + r_v\dot{v}, \quad r_u = \frac{\partial r}{\partial u}, \quad r_v = \frac{\partial r}{\partial v}. \tag{4}$$

如果我们所讨论的点是非异的, 则向量 $r_u = (x_u, y_u, z_u)$ 和 $r_v = (x_v, y_v, z_v)$ 为线性无关. 由 (4) 便得到结论说, 曲面的任意切向量是向量 r_u 和 r_v 的线性组合. 因此, 在所给非异点处切于此曲面的向量组成了以 r_u, r_v 为基底的二维空间, 并称其为这个点处的切平面. 在这组基下速度向量的坐标为 $\dot{u}, \dot{v}$.

假定曲面由方程 $F(x, y, z) = 0$ 给出, 且曲线 $r = r(t) = (x(t), y(t), z(t))$ 在此曲面上, 即 $F(x(t), y(t), z(t)) = 0$. 于是 $0 = \dfrac{d}{dt}F(x(t), y(t), z(t))$. 由此有

$$0 = F_x\dot{x} + F_y\dot{y} + F_z\dot{z}. \tag{5}$$

如果在所给点上 grad $F = (F_x, F_y, F_z) \neq 0$, 则对于切向量坐标 $(\dot{x}, \dot{y}, \dot{z})$ 的方程 (5) 便给出了切平面 (这个切平面方程具形式 $Ax + By + Cz = 0$, 其中 A, B, C 为导数 F_x, F_y, F_z 在已知 (非异) 点上所取的值.)

现设在具坐标 $(x^1, \cdots, x^n)$ 的 n 维空间 $\mathbb{R}^n$ 中给出了一个 k 维曲面 (以参数形式):

$$\begin{aligned} &x^1 = x^1(z^1, \cdots, z^k), \\ &\cdots\cdots\cdots\cdots \\ &x^n = x^n(z^1, \cdots, z^k). \end{aligned} \tag{6}$$

于是在曲面 (6) 上的曲线 $z^j = z^j(t), j = 1, \cdots, k$ 的速度向量有形式

$$v = (\dot{x}^1, \cdots, \dot{x}^n) = \dot{z}^1 f_1 + \cdots + \dot{z}^k f_k,$$

其中 $f_1, \cdots, f_k$ 为

$$f_j = \left(\frac{\partial x^1}{\partial z^j}, \cdots, \frac{\partial x^n}{\partial z^j}\right), \quad j = 1, \cdots, k. \tag{7}$$

向量 f_j 构成了在所给 (非异) 点上切于此平面的 k 维平面的基底. 我们看到了, 如果以坐标 $z^1, \cdots, z^k$ 给出了曲线, 则它的切向量在基 (f_j) 下具坐标 $(\dot{z}^1, \cdots, \dot{z}^k)$.

如果 k 维曲面由 $n-k$ 方程组

$$\begin{aligned} &F_1(x^1, \cdots, x^n) = 0, \\ &\cdots\cdots\cdots\cdots \\ &F_{n-k}(x^1, \cdots, x^n) = 0 \end{aligned} \tag{8}$$

给出, 则对切于此曲面的向量的坐标 $(\dot{x}^1, \cdots, \dot{x}^n)$ 满足线性方程组:

$$\sum_{l=1}^{n} \frac{\partial F_j}{\partial x^l} \dot{x}^l = 0, \quad j = 1, \cdots, n-k. \tag{9}$$

在非异点, 方程组 (9) 的秩等于 $n-k$, 这是因为方程组 (9) 定义了 $\mathbb{R}^n$ 中的一个 k 维平面. 这便是在所给点上曲面的切平面.

注 我们看到, 在所有情形下非异性条件的几何意义表现为: 在所给点上切于曲面的平面所具有的维数等于曲面的维数, 即对参数给出曲面时的参数个数 k. 如果曲面由方程组给出, 则这个数等于空间的维数减去方程个数.

3. 曲面上的度量

假设曲面 (或它的一部分) 由参数给出:$r = r(u,v), r = (x,y,z), u, v$ 为曲面上的坐标. 我们来讨论总是认定为非异的那些点: 矩阵 $A = \begin{pmatrix} x_u & y_u & z_u \\ x_v & y_v & z_v \end{pmatrix}$ 在这点的秩等于 2(其中 $x_u = \dfrac{\partial x}{\partial u}, x_v = \dfrac{\partial x}{\partial v}, \cdots$).

我们在前面是如何定义黎曼度量概念的?

设给出了曲线 $u = u(t), v = v(t)$. 它的长具有形式 $l = \int |v_t| dt, v_t = (\dot{u}, \dot{v})$. 这里的 $v_t = (\dot{u}, \dot{v})$ 是在坐标 (u,v) 下的速度向量, 而 $|v_t|^2 = g_{ij}\dot{x}^i\dot{x}^j$, 其中 $x^1 = u, x^2 = v, g_{ij} = g_{ij}(u,v)$. 我们曾称函数组 g_{ij} 为 (在坐标 u, v 下) 黎曼度量. 这组函数确定了曲线的长度, 还有两条曲线在它们相交处的夹角.

现在该如何定义曲线的长?$g_{ij}(u,v)(u = x^1, v = x^2)$ 又是什么?

我们注意, 曲线 $u = u(t), v = v(t)$ 是通过曲面上的坐标 (u,v) 写出来的, 而曲面本身则是在三维欧氏空间 (x,y,z) 中, 其中 $r = r(u,v), r = (x,y,z)$.

自然地, 我们称曲面上曲线 $u(t), v(t)$ 在三维欧氏空间中的长为它的长.

把曲线写成下面形状:

$$\begin{aligned} x &= x(u(t), \quad v(t)) = x(t), \\ y &= y(u(t), \quad v(t)) = y(t), \\ z &= z(u(t), \quad v(t)) = z(t). \end{aligned} \tag{10}$$

对三维欧氏空间中曲线的长度按定义我们有

$$l = \int_a^b \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}\,dt. \tag{11}$$

因为 $\dot{x} = x_u\dot{u} + x_v\dot{v}$, 故而我们得到

$$\dot{x}^2 + \dot{y}^2 + \dot{z}^2 = E\dot{u}^2 + 2F\dot{u}\dot{v} + G\dot{v}^2 = g_{ij}\dot{x}^j\dot{x}^i,$$

其中 $E = g_{11}, F = g_{12} = g_{21}, G = g_{22}, u = x^1, v = x^2$ 以及

$$\begin{aligned} g_{11} &= E = x_u x_u + y_u y_u + z_u z_u, \\ g_{12} &= F = x_u x_v + y_u y_v + z_u z_v, \\ g_{22} &= G = x_v x_v + y_v y_v + z_v z_v. \end{aligned} \tag{12}$$

如果令

$$r_u = x_u e_1 + y_u e_2 + z_u e_3, \quad r_v = x_v e_1 + y_v e_2 + z_v e_3,$$

则

$$g_{ij} = \langle r_{x^i}, r_{x^j}\rangle, \quad x^1 = u, \quad x^2 = v. \tag{13}$$

于是, 在曲面上的坐标下定义出了函数 $g_{ij}(u, v)$.

通常称表达式 $g_{ij}dx^i dx^j = E(du)^2 + 2F(dudv) + G(dv)^2$ 为曲面的第一基本型(或第一二次型, 或在曲面上诱导的黎曼度量).

如果曲面以 $F(x, y, z) = 0$ 的形式给出, 则曲面上的黎曼度量 (第一基本型) 就只是二次型 $dx^2 + dy^2 + dz^2$, 它满足条件 $F(x, y, z) = 0$; 由它得出

$$F_x dx + F_y dy + F_z dz = 0.$$

如果在所考虑的曲面的点上 $F_z \neq 0$, 则 $dz = -\left(\frac{F_x}{F_z}\right)dx - \left(\frac{F_y}{F_z}\right)dy$. 因此在曲面 $F(x, y, z) = 0, x = u, y = v$ 上有

$$dx^2 + dy^2 + dz^2 = dx^2 + dy^2 + \left(\frac{F_x}{F_z}dx + \frac{F_y}{F_z}dy\right)^2.$$

于是,

$$g_{11} = E = 1 + \frac{F_x^2}{F_z^2}, \quad g_{22} = G = 1 + \frac{F_y^2}{F_z^2}, \quad g_{12} = F = \frac{F_x F_y}{F_z^2}, \tag{14}$$

这里 $u = x^1 = x, v = x^2 = y$.

如果曲面的形式为 $z = f(x, y)$, 则在坐标 $u = x, v = y$ 下有

$$g_{11} = 1 + f_x^2, \quad g_{12} = f_x f_y, \quad g_{22} = 1 + f_y^2. \tag{15}$$

因此，曲面上的黎曼度量在这里就像是在这个曲面上以坐标 (u,v) 来计算曲线长度的一种方法．曲面自身是在三维欧氏空间之中，而这只不过是在三维欧几里得的意义下这一条曲线的长度问题．

在曲面上的欧氏坐标 (x,y,z) 以函数形式给出:$x=x(u,v),y=y(u,v),z=z(u,v)$. 由定义我们有

$$\begin{aligned} &dx^2+dy^2+dz^2=g_{ij}dx^idx^j, \quad x^1=u, \quad x^2=v; \\ &g_{11}=E, \quad g_{12}=F, \quad g_{22}=G. \end{aligned} \tag{16}$$

在某些情形会发生度量 $g_{ij}(x^1,x^2)$ 在曲面自身就是二维欧氏的 (参看 §3). 这表示在曲面上存在一对函数 $\bar{u}(x^1,x^2),\bar{v}(x^1,x^2)$, 使得

$$(d\bar{u})^2+(d\bar{v})^2=g_{ij}dx^idx^j. \tag{17}$$

例 7.5. 欧几里得柱面的度量: 柱面的方程为 $f(x,y)=0$(不含 z). 欧氏坐标为 z 及平面曲线 $f(x,y)=0$ 的自然参数 $l:\bar{u}=z,\bar{v}=l$. 我们有

$$dx^2+dy^2+dz^2|_{f(x,y)=0}=dz^2+dl^2.$$

假定曲面由 $F(x,y,z)=0$ 给出且 grad $F\neq 0$(曲面的非异点). 于是在曲面上有 $dF=F_xdx+F_ydy+F_zdz=0$. 如果 ξ 为曲面的切向量, $\xi=\xi^1e_1+\xi^2e_2+\xi^3e_3$, 则 $F_x\xi^1+F_y\xi^2+F_z\xi^3=0$, 或说 $\xi\perp$ grad F. 由此我们有结论: 向量 grad F(在非异点) 为曲面 $F(x,y,z)=0$ 的法方向.

在曲面的参数表示下:$r=r(u,v),r=(x,y,z)$, 我们有两个向量

$$\xi=r_u=r_{x^1}=x_ue_1+y_ue_2+z_ue_3,$$

$$\eta=r_v=r_{x^2}=x_ve_1+y_ve_2+z_ve_3.$$

这两个向量都切于曲面. 如果它们线性无关 (即非异点), 则它们的向量积 $[\xi,\eta]=[r_u,r_v]$ 正交于平面 (r_u,r_v), 于是正交于此曲面.

有了黎曼度量, 我们便能在曲面上度量任意曲线 $u=u(t),v=v(t)$ 的长度和两条曲线在它们的相交处的夹角. 在曲线 $(u_1(t),v_1(t))$ 和 $(u_2(t),v_2(t))$ 的交点 (u_0,v_0) 处各自的速度向量 $\eta_1=(\dot{u}_1,\dot{v}_1)$ 和 $\eta_2=(\dot{u}_2,\dot{v}_2)$ 之间的内积由公式

$$\langle\eta_1,\eta_2\rangle=g_{ij}\eta_1^i\eta_2^j \ (\eta_k^1=\dot{u}_k, \quad \eta_k^2=\dot{v}_k, \quad k=1,2)$$

给出，而它们之间的角 $\varphi(0\leqslant\varphi\leqslant\pi)$ 的公式是

$$\cos\ \varphi=\frac{\langle\eta_1,\eta_2\rangle}{|\eta_1||\eta_2|} \ (|\eta_1|^2=\langle\eta_1,\eta_1\rangle,|\eta_2|^2=\langle\eta_2,\eta_2\rangle).$$

现在我们推导在 n 维欧氏空间中 k 维曲面这种一般情形的公式. 在第 1 小节中所讨论的使我们可以限于以参数形式给出的曲面: $x^i = x^i(z^1, \cdots, z^k), i = 1, \cdots, n$. 在空间中的曲面

$$x^i(t) = x^i(z^1(t), \cdots, z^k(t)); \quad i = 1, \cdots, n, \tag{18}$$

上的曲线 $z^j = z^j(t), j = 1, \cdots, k$, 其长度的计算公式是

$$l = \int_a^b |\dot{x}| dt = \int_a^b \sqrt{\sum_{i=1}^n (\dot{x}^i)^2}\, dt = \int_a^b \sqrt{\sum_{q=1}^n \left(\frac{\partial x^q}{\partial z^i}\frac{\partial x^q}{\partial z^j}\right)\dot{z}^i\dot{z}^j}\, dt = \int_a^b \sqrt{g_{ij}\dot{z}^i\dot{z}^j}\, dt, \tag{19}$$

$$g_{ij}(z^1, \cdots, z^k) = \sum_{q=1}^n \frac{\partial x^q}{\partial z^i}\frac{\partial x^q}{\partial z^j}. \tag{20}$$

显然, $(dl)^2 = g_{ij}dz^i dz^j$.

因此, 空间的度量决定了置于其中的任意曲面上的度量; 赋予的这个度量一般来说是非欧几里得的. 称度量 (20) 为在曲面上的*诱导度量*.

对于超曲面 $F(x^1, \cdots, x^n) = 0, \text{grad } F \neq 0$ 的情形 (设 $\partial F/\partial x^n \neq 0$), 度量 (20) 有形式

$$g_{ij} = \delta_{ij} + \frac{F_{x^i}F_{x^j}}{F_{x^n}^2},$$

其中 $F_{x^q} = \dfrac{\partial F}{\partial x^q}$.

4. 曲面的面积

如果我们有具坐标 x, y 的欧氏平面及平面上的区域 U, 则区域 U 的面积可用下面的二重积分 $\sigma(U)$ 来量度:

$$\sigma(U) = \iint_U dxdy.$$

如果我们作变量变换 (一一对应的)

$$x = x(u, v), \quad y = y(u, v), \tag{21}$$

则我们有公式

$$\sigma(U) = \iint_V |x_u y_v - x_v y_u| dudv, \tag{22}$$

其中 V 为 (u, v) 平面中的区域, 它对应于 (x, y) 平面中的区域 U. 因此我们有

$$\sigma(U) = \iint_V |J| dudv,$$

其中 J 是变量变换 (21) 的雅可比, $J = x_u y_v - y_u x_v$. 问题是: 如果知道了在曲面 $r = r(u,v), r = r(x,y,z)$ 上的黎曼度量, 那么如何计算此曲面上区域 U 的面积? 这时的黎曼度量假定为

$$dl^2 = g_{ij}dx^i dx^j, \quad x^1 = u, \quad x^2 = v. \tag{23}$$

考虑矩阵 (g_{ij}) 的行列式:

$$g = \det(g_{ij}) = g_{11}g_{22} - g_{12}^2 = EG - F^2 > 0. \tag{24}$$

定义 7.4. 称

$$\sigma(U) = \iint_U \sqrt{g}dudv \tag{25}$$

为曲面 $r = r(u,v), r = (x,y,z)$ 上区域 U 的面积, 其中 U 作为 (u,v) 平面中由参数给出的区域.

称 $d\sigma = \sqrt{g}dudv$ 为在黎曼度量 (g_{ij}) 下曲面的面积微分 (或面积元).

如此定义曲面上区域的面积的基础是什么? 如果在点 (u,v) 的切向量的内积由矩阵 (g_{ij}) 给出, 那么为什么面积元应该选取 $\sqrt{g}dudv$ 这种形式?

为了了解这些问题, 我们考虑欧氏平面的一对向量 ξ, η 及平行四边形 $\lambda\xi + \mu\eta$, $0\leqslant\lambda\leqslant1, 0\leqslant\mu\leqslant1$. 此平行四边形的面积等于 $\sigma = |\xi^1\eta^2 - \xi^2\eta^1| = |\det A|$, 其中 $A = \begin{pmatrix} \xi^1 & \xi^2 \\ \eta^1 & \eta^2 \end{pmatrix}$, 即由向量 $\xi = \xi^1 e_1 + \xi^2 e_2$ 和 $\eta = \eta^1 e_1 + \eta^2 e_2$ 对于法正交基 e_1, e_2 的分量构成的矩阵.

现在提出另一个问题: 假设给平面以基向量 $\bar{e}_1, \bar{e}_2$. 又假设基向量的内积有矩阵

$$\langle \bar{e}_i, \bar{e}_j \rangle = g_{ij}, \quad i = 1,2, \quad j = 1,2. \tag{26}$$

由向量 $\bar{e}_1$ 和 $\bar{e}_2$ 张成的平行四边形的面积是什么? 此平行四边形的点为 $\lambda\bar{e}_1 + \mu\bar{e}_2, 0\leqslant\lambda\leqslant1, 0\leqslant\mu\leqslant1$. 我们假定矩阵 (g_{ij}) 为正定二次型

$$g_{ij}\xi^i\xi^j > 0$$

的矩阵.

引理 7.2. 由向量 $\bar{e}_1$ 和 $\bar{e}_2$ 张成的平行四边形的面积等于 $\sqrt{g}$, 其中

$$g = \det(g_{ij}) = g_{11}g_{22} - g_{12}^2.$$

证明 二次型 g_{ij} 可由线性变换 A 化为对角形:$g'_{ij} = \delta_{ij}$. 准确地就是: 求出向量 e_1, e_2 :

$$\bar{e}_1 = a_{11}e_1 + a_{12}e_2, \quad \bar{e}_2 = a_{21}e_1 + a_{22}e_2, \tag{27}$$

使得 $\langle e_i, e_j\rangle = g'_{ij} = \delta_{ij}$ (即 $|e_i|^2=1, e_1 \perp e_2$). 由 (27) 得到

$$g_{11} = \langle \bar{e}_1, e_1\rangle = a_{11}^2 + a_{12}^2,$$
$$g_{12} = g_{21} = \langle \bar{e}_1, \bar{e}_2\rangle = a_{11}a_{21} + a_{22}a_{12},$$
$$g_{22} = \langle \bar{e}_2, \bar{e}_2\rangle = a_{21}^2 + a_{22}^2.$$

用矩阵表示即

$$(g_{ij}) = G = AA^{\mathrm{T}}, \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}. \tag{28}$$

那么由向量 $\bar{e}_1, \bar{e}_2$ 张成的平行四边形的面积是什么呢? 因为基 (e_1, e_2) 为法正交且向量 $\bar{e}_1, \bar{e}_2$ 有 (27) 的形状, 故由 $\bar{e}_1, \bar{e}_2$ 张成的平行四边形面积等于 $|\det A|$. 但 $\det A^{\mathrm{T}} = \det A$, 于是由 (28) 得到

$$g = \det(g_{ij}) = (\det A)^2, \quad |\det A| = \sqrt{g}. \tag{29}$$

引理得证. □

我们现在不加证明地回忆在二重积分基本概念和与区域面积相关的概念中的一些基本思想.

考虑有坐标 $x^1 = u, x^2 = v$ 的平面中的区域 U, 它由一些分段光滑的曲线 Γ 围成. 我们将平面分割成以 Δu 和 Δv 为边的小矩形 (我们假定 Δu 和 Δv 趋向于零). 显然, 区域 U 的面积大于 (或等于) 所有在内部的这些矩形区域的面积和.

定义 7.5. 称所有内部的这些矩形的面积和在 $\Delta u \to 0, \Delta v \to 0$ 时的极限为区域 U 的面积; 这时我们假定了这个极限是存在的.

现在另外设有一个二元连续函数 $f(u,v)$. 我们来定义这个函数在区域 U 上的积分.

考虑区域内部所有以 Δu 和 Δv 为边的矩形构成的矩形网格. 对于矩形 S_α 我们考虑函数在矩形中心的值 $f(u_\alpha, v_\alpha)$. 作积分和

$$S(f,U) = \sum_\alpha f(u_\alpha, v_\alpha)\Delta u \Delta v,$$

其中是对所有内部的矩形取和.

定义 7.6. 如果当 $\Delta u \to 0, \Delta v \to 0$ 时, 和 $S(f,U)$ 的极限存在, 我们称这个极限为函数 $f(u,v)$ 在此区域上的二重积分, 并记为 $\iint\limits_U f(u,v)dudv$.

特别, 当 $f(u,v) \equiv 1$, 且 u, v 为欧氏坐标时, 积分 $\iint\limits_U dudv$ 就是区域 U 的面积.

回忆前面我们在坐标 $u=x^1, v=x^2$ 的平面上定义了区域 U 的面积, 这时的度量有形式 $dl^2=g_{11}du^2+2g_{12}dudv+g_{22}dv^2$, 而面积为积分

$$\sigma(U)=\iint\limits_U \sqrt{g}dudv.$$

为什么? 如果 $\Delta u, \Delta v$ 是小的, 则中心在点 (u_α, v_α), 边为 Δu 和 Δv 的小平行四边形的面积 $S_\alpha \approx \Delta u\Delta v\sqrt{g}$, 这由证明过的上述论断得到. 这里的 $\sqrt{g}=\sqrt{g_{11}g_{12}-g_{12}^2}$, 其中 g_{ij} 在点 (u_α, v_α) 取值. 对于小的 $\Delta u, \Delta v$, 我们有

$$\sum_\alpha S_\alpha \approx \sum_\alpha \sqrt{g(u_\alpha, v_\alpha)}\Delta u\Delta v.$$

在 $\Delta u\to 0, \Delta v\to 0$ 时这个和的极限为积分 $\sigma(U)=\iint\limits_U \sqrt{g}dudv$. 最后我们推导出在各种表示下的曲面面积公式.

定理 7.2. 1. 设曲面以图像形式 $z=f(x,y)$ 给出, 又设曲面上所给区域 U 投影到 (x,y) 平面上为区域 V. 于是有等式

$$\sigma(U)=\iint\limits_V \sqrt{1+f_x^2+f_y^2}\, dxdy. \tag{30}$$

2. 如果曲面由 $F(x,y,z)=0$ 给出, 而曲面上的区域 U 一一地投射为 (x,y) 平面上的区域 V. 于是我们有公式

$$\sigma(U)=\iint\limits_V \frac{|\text{grad } F|}{|F_z|}dxdy, \tag{31}$$

其中 $F_z=\dfrac{\partial F}{\partial z}$ 在 (x,y,z) 属于区域 U 时不为零.

3. 如果曲面以参数形式给出:$r=r(u,v)$, 则有公式

$$\sigma(U)=\iint\limits_V |[r_u, r_v]|dudv, \tag{32}$$

其中 V 为 (u,v) 平面上对应于区域 U 的区域, $[r_u, r_v]$ 为向量积.

证明 1. 回忆对于曲面 $z=f(x,y)$ 我们有

$$g_{11}=1+f_x^2, \quad g_{12}=f_xf_y, \quad g_{22}=1+f_y^2, \quad u=x, \quad v=y.$$

于是 $\sqrt{g}=\sqrt{g_{11}g_{22}-g_{12}^2}=\sqrt{1+f_x^2+f_y^2}$. 由定义 7.4 给出的曲面面积便得到此重要论断.

2. 对于 $F_z \neq 0$ 的 $F(x,y,z)=0$, 有 $x^1=u=x, x^2=v=y, g_{11}=1+\frac{F_x^2}{F_z^2}, g_{12}=\frac{F_xF_y}{F_z^2}, g_{22}=1+\frac{F_y^2}{F_z^2}$. 像在第 1 小节中那样可验证

$$\sqrt{g}=\sqrt{1+\frac{F_x^2}{F_z^2}+\frac{F_y^2}{F_z^2}}=\frac{|\text{grad } F|}{|F_z|}.$$

3. 回忆对于 $r=r(u,v), x^1=u, x^2=v$, 我们有 $g_{ij}=\langle r_{x^i}, r_{x^j}\rangle$. 因此对由向量 r_u 和 r_v 张成面积等于 $|[r_u,r_v]|$ 的平行四边形, 我们有

$$g=g_{11}g_{22}-g_{12}^2=|[r_u,r_v]|^2.$$

于是由定义 7.4 得

$$\sigma(U)=\iint_V |[r_u,r_v]|dudv.$$

定理得证. □

这样, 我们相信面积确实像长度一样, 能够被对在每点的切向量给出的内积 (g_{ij}) 所确定.

习题

7.1. 三维空间中的环面 T^2 是围绕直线旋转形成的曲面, 而这条直线处在此圆所在的平面中. 写出此环面的参数方程和在此环面上的诱导度量.

7.2. 计算出旋转椭球面的第一基本形式; 椭球面方程为

$$\frac{x^2}{a^2}+\frac{y^2+z^2}{b^2}=1.$$

7.3. 求在旋转面

$$r(u,\varphi)=(\rho(u)\cos\ \varphi, \rho(u)\sin\ \varphi, z(u))$$

上的诱导度量. 验证它的经线 ($\varphi=$ 常数) 和纬线 ($u=$ 常数) 构成正交网格. 求平分经纬线之间夹角的曲线.

7.4. 求球面上的曲线, 它以已知角 α 交于经线 (斜驶线). 求斜驶线的长.

7.5. 设 $F(x,y,z)$ 为光滑的齐次函数 (即 $F(cx,cy,cz)=c^nF(x,y,z)$). 证明: 在锥面 $F(x,y,z)=0$ 上的度量在坐标原点之外是欧几里得的.

§8. 第二基本型

1. 欧氏空间中曲面上曲线的曲率

设在三维欧氏空间中给出了曲面，而 (x_0, y_0, z_0) 为它上面的非异点．我们首先假设 z 轴垂直于曲面在点 (x_0, y_0, z_0) 的切平面．于是曲面在点 (x_0, y_0, z_0) 附近由方程 $z = f(x, y)$ 局部地给出，其中 $z_0 = f(x_0, y_0)$，且

$$\left.\frac{\partial f}{\partial x}\right|_{\substack{x=x_0\\y=y_0}} = \left.\frac{\partial f}{\partial y}\right|_{\substack{x=x_0\\y=y_0}} = 0, \quad \text{即 grad } f\Big|_{\substack{x=x_0\\y=y_0}} = 0. \tag{1}$$

考虑函数 $z = f(x, y)$ 的二阶微分，即 $d^2 f = f_{xx}dx^2 + 2f_{xy}dxdy + f_{yy}dy^2$，并作矩阵 $(a_{ij}), a_{ij} = f_{x^i x^j}$，其中 $x^1 = x, x^2 = y$ (称此矩阵为黑塞矩阵)．考虑在点 (x_0, y_0, z_0) 满足 grad $f = 0$ 的这个二次型的矩阵．

定义 8.1． 设曲面 $z = f(x, y)$ 在点 (x_0, y_0, z_0) 有 grad $f = 0$．称矩阵 (a_{ij}) 的特征值为此曲面在这点的*主曲率*．称矩阵 (a_{ij}) 的行列式为在这点的*高斯曲率*．称矩阵的迹为在此点的*平均曲率*(或*中曲率*)：迹 $a_{11} + a_{22} = k_1 + k_2$，$k_1$ 和 k_2 为特征值，高斯曲率 $K = k_1 k_2 = \det(a_{ij}) = a_{11}a_{22} - a_{12}a_{21}$．

我们将在第四章看到，曲面的高斯曲率只与这个曲面内在的度量性质有关．

到现在我们是在一个与所讨论的点有关的特殊的坐标下定义了曲率的概念：要求 z 轴垂直于曲面，而轴 x 和 y 在这点切于它，或者说，局部地 $z = f(x, y)$ 的 grad f 在点 (x_0, y_0, z_0) 为 0．为了在任意坐标下定义这些量，我们转向曲面上曲线的曲率理论．

设曲面以参数形式给出：

$$r = r(u, v). \tag{2}$$

于是在每个非异点有 $[r_u, r_v] = |[r_u, r_v]|m$，其中 m 为曲面的单位法向量，$|m| = 1$．考虑曲面上曲线 $r = r(u(t), v(t))$．我们有 $\dot{r} = r_u\dot{u} + r_v\dot{v}, \ddot{r} = (r_{uu}\dot{u}^2 + 2r_{uv}\dot{u}\dot{v} + r_{vv}\dot{v}^2) + (r_u\ddot{u} + r_v\ddot{v})$．因为 $r_u \perp m$ 和 $r_v \perp m$，我们得到

$$\begin{aligned}\langle \ddot{r}, m\rangle &= \langle r_{uu}, m\rangle\dot{u}^2 + 2\langle r_{uv}, m\rangle\dot{u}\dot{v} + \langle r_{vv}, m\rangle\dot{v}^2\\ &= b_{11}\dot{u}^2 + 2b_{12}\dot{u}\dot{v} + b_{22}\dot{v}^2.\end{aligned} \tag{3}$$

结论 加速度的法向投影 $\langle \ddot{r}, m\rangle$ 是在局部坐标 $u = x^1, v = x^2$ 下速度向量 $(\dot{u}, \dot{v})$ 的二次型．

令 $b_{11} = L, b_{12} = M, b_{22} = N$．我们有

$$\langle \ddot{r}, m\rangle dt^2 = b_{ij}dx^i dx^j = Ldu^2 + 2Mdudv + Ndv^2.$$

定义 8.2． 称表达式 $\langle \ddot{r}, m\rangle dt^2$ 为曲面 (2) 的*第二基本型*(*第二二次型*)．

设曲线 $u(t), v(t)$ 参照于自然参数 $t = l$. 依照对曲线 $r = r(t) = r(u(t), v(t))$ 的弗莱纳公式, 我们有

$$\ddot{r} = \frac{d^2 r}{dl^2} = kn,$$

其中 n 为此曲线的主法线, k 为曲线的曲率. 因此 $\langle \ddot{r}, m\rangle = k\langle n, m\rangle = k\cos\ \theta$, 其中 θ 为 m 与 n 间的角. 由此得到

$$k\cos\ \theta dl^2 = \langle \ddot{r}, m\rangle dl^2 = Ldu^2 + 2M\,dudv + Ndv^2 = b_{ij}dx^i dx^j,$$

$$x^1 = u, \quad x^2 = v,$$

其中 $dl^2 = g_{ij}dx^i dx^j = Edu^2 + 2Fdudv + Gdv^2$.

结论

$$k\cos\ \theta = \frac{b_{ij}dx^i dx^j}{g_{ij}dx^i dx^j}, \quad x^1 = u, \quad x^2 = v. \tag{4}$$

(应该记住, 在相似的记号下 dx^i 并不是一个独立的微分, 而应理解为 $dx^i = \dot{x}^i dt$.) 进而证明了

定理 8.1. 三维空间中曲面上的曲线的曲率乘以曲面法线与曲线主法线间夹角的余弦, 等于在曲线切向量上第二基本型与第一基本型的比值.

推论 如果曲线为曲面与曲面在这点的法平面的交线, 则 $\cos\ \theta = 1$, 从而

$$k = \frac{b_{ij}dx^i dx^j}{g_{ij}dx^i dx^j}. \tag{5}$$

2. 二次型偶对的不变量

现在, 在曲面的每一点都有了一对二次型

1)　$dl^2 = g_{ij}dx^i dx^j$,　(6)

2)　$\langle \ddot{r}, m\rangle dt^2 = b_{ij}dx^i dx^j$.　(7)

并且形式 dl^2 为正定.

在代数中我们知道对一个二次型的偶对有什么样的不变量?

考虑平面上的一对二次型, 它们中的一个是正定的. 设这两个二次型的矩阵分别为

$$G = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}, \quad Q = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \tag{8}$$

$(g_{21} = g_{12}, b_{21} = b_{12})$. 作方程

$$\det(Q - \lambda G) = 0, \tag{9}$$

或

$$(b_{11} - \lambda g_{11})(b_{22} - \lambda g_{22}) - (b_{12} - \lambda g_{12})^2 = 0.$$

称这个方程的根 λ_1, λ_2 为这个二次型偶对的特征值.

解线性方程组

$$\left.\begin{aligned}(b_{11}-\lambda_i g_{11})\xi_i^1+(b_{12}-\lambda_i g_{12})\xi_i^2=0,\\(b_{12}-\lambda_i g_{12})\xi_i^1+(b_{22}-\lambda_i g_{22})\xi_i^2=0,\end{aligned}\right\}\tag{10}$$

其中 ξ_i^1,ξ_i^2 为未知的, $i=1,2$.

如果 λ_1,λ_2 为特征值, 则方程组 (10) 有非平凡解

$$f_1=(\xi_1^1,\xi_1^2)\quad 和\quad f_2=(\xi_2^1,\xi_2^2).$$

向量 f_1,f_2 的方向被称做这个二次型偶对的*主方向*; 其中 f_1 对应于 λ_1, f_2 对应于 λ_2.

像以前那样, 以 $g_{ij},i,j=1,2$ 代表平面的基向量的内积 $\langle e_i,e_j\rangle$(于是形式 g_{ij} 给出了黎曼度量).

引理 8.1. 如果二次型偶对的特征值互不相同, 则主方向正交.

我们有两个主方向 f_1,f_2,

$$f_1=\xi_1^1e_1+\xi_1^2e_2,\quad f_2=\xi_2^1e_1+\xi_2^2e_2.$$

按定义, 正交性的意思是

$$\langle f_1,f_2\rangle=g_{ij}\xi_1^i\xi_2^j=0.$$

证明 我们选取平面的一对向量 d_1,d_2, 使 $\langle d_i,d_j\rangle=\delta_{ij}$. 由于矩阵 (g_{ij}) 对应的二次型是正定的, 这样的 d_1,d_2 存在. 因为可经线性变换使此矩阵变为平方和. 我们现在在新基底 d_1,d_2 下考察第二基本型. 设

$$e_i=a_i^jd_j,\quad A=(a_i^j).\tag{11}$$

在新基底 d_1,d_2 下的第一和第二基本型分别记为 $\bar G,\bar Q$, 则

1) $\bar G=(\langle d_i,d_j\rangle)=(\delta_{ij})$ 或 $\bar G=\begin{pmatrix}1&0\\0&1\end{pmatrix}$, 故而 $G=A^{\mathrm T}A$;

2) $Q=A^{\mathrm T}\bar QA$.

因为 $G=A^{\mathrm T}A,Q=A^{\mathrm T}\bar QA$, 故

$$Q-\lambda G=A^{\mathrm T}(\bar Q-\lambda\cdot 1)A,$$

$$\det(Q-\lambda G)=(\det\ A)^2\det(\bar Q-\lambda\cdot 1).\tag{12}$$

注意, $\det A=\det A^{\mathrm T}=\sqrt{g}=\sqrt{\det G}\neq 0$. 所以方程 (9) 等价于方程

$$\det(\bar Q-\lambda\cdot 1)=0.\tag{13}$$

在基底 d_1,d_2 下的内积是欧几里得的. 它给出了单位矩阵 $\bar G=1=(\delta_{ij})$. 由代数课程知道, 二次型 $\bar Q$ 可经旋转化为具对角矩阵的二次型, 而它的特征向量 f_1,f_2 在通常欧氏意义下为正交. 引理得证. □

这个引理是关于经旋转化欧氏平面中二次型为对角型的定理的另一种形式.

3. 第二基本型的性质

我们现在回到三维欧氏空间中曲面的第一和第二基本型:

$$g_{ij}dx^idx^j = dl^2 \tag{14}$$

$$b_{ij}dx^idx^j = \langle \ddot{r}, m\rangle dl^2, \quad x^1 = u, \quad x^2 = v. \tag{15}$$

这两个二次型的比值是法平面交线的曲率.

定义 8.3. 称这个二次型偶对的特征值为在所给点上曲面的*主曲率*. 称这两个主曲率的积为曲面的*高斯曲率*, 而它们的和称为曲面的*平均曲率* ①(或*中曲率*).

例 8.1. 设曲面由形式 $z = f(x, y)$ 给出, 并设在所讨论的点 (x_0, y_0) 上有 $f_x = f_y = 0$ (即在此点 z 轴垂直于曲面的切平面). 设 $x = u, y = v, z = f(u, v)$. 对第一和第二基本型 (在所给点 (x_0, y_0) 上) 我们有

$$\begin{aligned}
&(\text{I}) \quad g_{11} = 1, \quad g_{22} = 1, \quad g_{12} = g_{21} = 0 \ (g_{ij} = \delta_{ij}),\\
&(\text{II}) \quad L = b_{11} = \langle r_{uu}, m\rangle = f_{xx}(x_0, y_0),\\
&\qquad M = b_{12} = \langle r_{uv}, m\rangle = f_{xy}(x_0, y_0),\\
&\qquad N = b_{22} = \langle r_{vv}, m\rangle = f_{yy}(x_0, y_0),
\end{aligned} \tag{16}$$

其中向量 m 同于 z 轴上的单位向量.

因此, 在所研究的点上第二基本型有形式

$$b_{ij}dx^idx^j = f_{x^ix^j}dx^idx^j = d^2f. \tag{17}$$

这时的高斯曲率便等于行列式

$$\det\begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy}\end{pmatrix}.$$

由方程 $(f_{xx} - \lambda)(f_{yy} - \lambda) - (f_{xy})^2 = 0$ 可求出特征值, 因为这时我们有 $g_{ij} = \delta_{ij}$.

在这个点上曲面的切平面平行于 (x, y) 平面. 而在此点的主方向可通过解下面方程得到:

$$(f_{x^ix^j} - \lambda_1\delta_{ij})\xi_1^j = 0, \quad i, j = 1, 2, \quad 解出 f_1, \tag{18}$$

$$(f_{x^ix^j} - \lambda_2\delta_{ij})\xi_2^j = 0, \quad i, j = 1, 2, \quad 解出 f_2, \tag{19}$$

由引理 8.1 知 $f_1 \perp f_2$ (这里像以前那样, $f_i = (\xi_i^1, \xi_i^2), i = 1, 2$), 故而我们可以以单位主方向向量为新的坐标 x', y', 只要进行 (x, y) 平面中的旋转即可. 为此只需 $\lambda_1 \neq \lambda_2$.

①常称主曲率的半和为平均曲率.

在新坐标 (x',y',z) 下，我们有

$$z=f(x(x',y'),y(x',y')),$$

其中 $x=x'\cos\varphi+y'\sin\varphi, y=-x'\sin\varphi+y'\cos\varphi, \varphi$ 为旋转角.

在新坐标下，第二基本型具有形式 (只在所考虑的点上)

$$\lambda_1(dx')^2+\lambda_2(dy')^2. \tag{20}$$

我们得到了在所给点处的法截线的曲率为

$$k=\frac{\lambda_1(dx')^2+\lambda_2(dy')^2}{(dx')^2+(dy')^2}. \tag{21}$$

在所给点处的法截线的切向量 e 为 $e=(\dot{x}',\dot{y}')$, 其中

$$dx'=\dot{x}'dt, \quad dy'=\dot{y}'dt.$$

因此我们有

$$\cos^2\alpha=\frac{(dx')^2}{(dx')^2+(dy')^2}, \quad \sin^2\alpha=\frac{(dy')^2}{(dx')^2+(dy')^2}, \tag{22}$$

其中 α 为 x 轴与法截线的切向量之间的夹角.

因此，在我们特别选取的坐标系下建立了等式 $k=\lambda_1\cos^2\alpha+\lambda_2\sin^2\alpha$. 称此公式为欧拉公式. 事实上，它在任意坐标系下都成立.

定理 8.2. 法截线的曲率为

$$k=\lambda_1\cos^2\alpha+\lambda_2\sin^2\alpha,$$

其中 λ_1,λ_2 为主曲率，α 为曲面上法截线的切线与对应于 λ_1 的主方向之间的夹角.

证明 我们在曲面由 $z=f(x,y)$ 给出并在所考虑的点 (x_0,y_0) 有 $f_x=f_y=0$ 的情况下证明欧拉公式. 但是，由于所要证明的结果不依赖于坐标的选取，故而我们总能在所考虑的点上选取相关的坐标使得 z 轴与曲面在这点正交，而 x 和 y 轴切于此曲面并相互正交 (甚至可以是主方向). 于是在已知点的邻域中曲面有形式

$$z=f(x,y), \quad f_x=f_y=0.$$

除此而外，如果 x,y 轴是主方向的话，在此点更有 $f_{xy}=f_{yx}=0$. 那么 $\lambda_1=f_{xx},\lambda_2=f_{yy}$. 因为我们已经在这样的坐标中推导出了欧拉公式，故定理得证. □

注 把 z 轴的方向变为相反时主曲率的符号也随之改变.

注意欧拉公式的直接推论: 如果 $\lambda_1>\lambda_2$, 则 λ_1 和 λ_2 是在所给点处最大和最小的法截线曲率(考虑符号).

我们来给出对第二基本型有用的公式.

如果曲面以 $z=f(x,y)$ 形式给出, 则第二基本型的系数为 (这里 $x=u, y=v$)

$$r_u=(1,0,f_x),\quad r_v=(0,1,f_y),\quad [r_u,r_v]=(-f_x,-f_y,1),$$

$$r_{uu}=(0,0,f_{xx}),\quad r_{uv}=(0,0,f_{xy}),\quad r_{vv}=(0,0,f_{yy}),$$

$$\begin{aligned} m&=\frac{[r_u,r_v]}{|[r_u,r_v]|}=\frac{(-f_x,-f_y,1)}{\sqrt{1+f_x^2+f_y^2}},\quad L=b_{11}=\frac{f_{xx}}{\sqrt{1+f_x^2+f_y^2}},\\ M&=b_{12}=b_{21}=\frac{f_{xy}}{\sqrt{1+f_x^2+f_y^2}},\quad N=b_{22}=\frac{f_{yy}}{\sqrt{1+f_x^2+f_y^2}}. \end{aligned} \tag{23}$$

由此得到

$$b_{ij}dx^i dx^j=\frac{1}{\sqrt{1+f_{x^1}^2+f_{x^2}^2}}(f_{x^ix^j}dx^idx^j), \tag{24}$$

$$x^1=u=x,\quad x^2=v=y.$$

回想我们对系数 g_{ij} 有公式 (7.15):

$$\begin{aligned} &g_{11}=1+f_x^2,\quad g_{12}=g_{21}=f_xf_y,\quad g_{22}=1+f_y^2,\\ &g=g_{11}g_{22}-g_{12}^2=1+f_x^2+f_y^2. \end{aligned} \tag{25}$$

我们有下面的定理:

定理 8.3. 曲面的高斯曲率为第二基本型和第一基本型的行列式的比值:

$$K=\frac{b_{11}b_{22}-b_{12}^2}{g_{11}g_{22}-g_{12}^2}. \tag{26}$$

特别地, 如果曲面以图像 $z=f(x,y)$ 的形式给出, 则有公式

$$K=\frac{f_{xx}f_{yy}-f_{xy}^2}{(1+f_x^2+f_y^2)^2}.$$

证明 特征值 λ_1 和 λ_2 由方程 (9)

$$\det(Q-\lambda G)=0$$

决定, 其中 $Q=\begin{pmatrix} L & M\\ M & N\end{pmatrix}$ 为第二基本型的矩阵, 而 $G=(g_{ij})$. 我们注意到 $Q-\lambda G=(b_{ij}-\lambda g_{ij})$ 且 $\det(Q-\lambda G)=(b_{11}-\lambda g_{11})(b_{22}-\lambda g_{22})-(b_{12}-\lambda g_{12})^2$. 矩阵 $G=(g_{ij})$ 正定, 从而非异. 因而

$$\det(Q-\lambda G)=\det G\cdot\det(G^{-1}Q-\lambda\cdot 1),$$

其中 $\det G \neq 0$. 解方程 $\det(G^{-1}Q - \lambda \cdot 1) = 0$, 从而可确定特征值 λ_1 和 λ_2. 回想代数中的结果: 矩阵的所有特征值的积等于它的行列式, 特别, $\det A = \lambda_1\lambda_2$ 是二阶矩阵的情形. 令 $A = G^{-1}Q$, 我们有

$$\lambda_1\lambda_2 = \det(G^{-1}Q) = \frac{\det Q}{\det G} = K. \tag{27}$$

由此知道高斯曲率等于第二基本型和第一基本型的矩阵的行列式的比值.

更进一步, 如果曲面由 $z = f(x,y)$ 给出, 则 b_{ij}, g_{ij} 由公式 (23) 和 (7.15) 给出. 计算行列式并作出比值便得到高斯曲率的公式. □

推论 如果曲面由图像 $z = f(x,y)$ 的形式给出, 则高斯曲率的符号与行列式 $f_{xx}f_{yy} - f_{xy}^2$ 的符号一致, 这是因为有

$$K = \frac{f_{xx}f_{yy} - f_{xy}^2}{(1 + f_x^2 + f_y^2)^2}.$$

例 8.2. 设曲面由 $z = f(x,y)$ 给出, 其中函数 $f(x,y)$ 满足拉普拉斯方程 $f_{xx} + f_{yy} = 0$. 因为 $f_{xx} = -f_{yy}$, 故 $f_{xx}f_{yy} - f_{xy}^2 \leqslant 0$. 于是高斯曲率处处有 $K < 0$, 其中两个导数 f_{xx}, f_{xy} 中至少有一个不为零.

我们要找出高斯曲率的几何意义. 对所给曲面上的点 (x_0, y_0, z_0), 我们选取法正交标架 (x, y, z), 其中 z 轴正交于曲面. 于是局部地曲面可写成 $z = f(x,y)$, 其中 f_x 和 f_y 在此点为 0.

在所给点我们有 $g_{ij} = \delta_{ij}$, 这是因为这时 $g_{11} = 1 + f_x^2, g_{12} = f_xf_y, g_{22} = 1 + f_y^2$; 另外 $L = b_{11} = f_{xx}, M = b_{12} = f_{xy}, N = b_{22} = f_{yy}$.

我们分析三种情形.

1. $K > 0, \lambda_1 > 0, \lambda_2 > 0$ (在点 $x = x_0, y = y_0$ 函数 $f(x,y)$ 取极小, 图 10,a).
2. $K > 0, \lambda_1 < 0, \lambda_2 < 0$ (在点 $x = x_0, y = y_0$ 函数 $f(x,y)$ 取极大, 图 10,b).
3. $K < 0$, 因此 $\lambda_1 > 0, \lambda_2 < 0$ 或者相反, 这是鞍点, 或 "交叉" 点 (图 10,c).

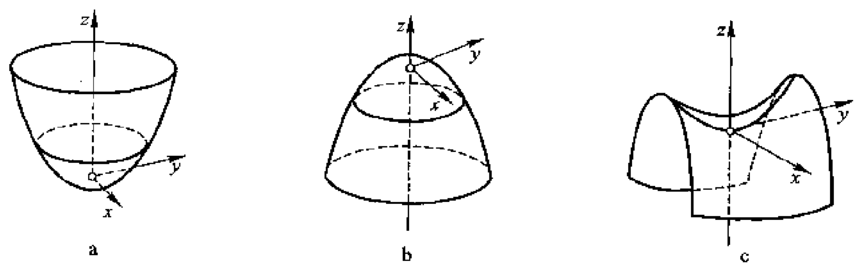


图 10

结论 在 $K > 0$ 时曲面局部地处于所考虑点处的切平面的一侧. 在 $K < 0$ 时曲面必定在任意靠近切点处与此点的切平面相交.

如果高斯曲率处处为正, 则此曲面是严格凸曲面.

习题

8.1. 求一曲面, 其所有法线交于一个点.

8.2. 计算下面旋转面的第二基本型:

$$r(u,\varphi) = (x(u), \rho(u)\cos\varphi, \rho(u)\sin\varphi), \rho(u) > 0.$$

8.3. 计算下面曲面的高斯曲率和平均曲率, 曲面方程为

$$z = f(x) + g(y).$$

8.4. 证明: 如果三维欧氏空间中的曲面其高斯和平均曲率恒同于零, 则此曲面为平面.

8.5. 证明: 曲面 $z = f(x,y)$ 上平均曲率 H 等于

$$H = \operatorname{div}\left(\frac{\operatorname{grad} f}{\sqrt{1+|\operatorname{grad} f|^2}}\right).$$

8.6. 设曲面 S 由一条具曲率 $k(l)$ 的已知曲线的切直线构成. 证明: 如果曲线在保持 $k(l)$ 不变情形下扭曲, 则相应的曲面保持了度量不变.

8.7. 如果曲面有形式

$$dl^2 = A^2du^2 + B^2dv^2, A = A(u,v), B = B(u,v),$$

则其高斯曲率为

$$K = -\frac{1}{AB}\left[\left(\frac{A_v}{B}\right)_v + \left(\frac{B_u}{A}\right)_u\right].$$

8.8. 证明: 具有零平均曲率的唯一的旋转曲面为平面和悬链面, 这里的悬链面是由曲线 $x = \dfrac{\operatorname{ch}(ay+b)}{a}$ 绕 y 轴旋转得到.

§9. 球面的度量

半径为 R, 球心在坐标原点的球面 $S^2 \subset \mathbb{R}^3$ 由方程

$$x^2 + y^2 + z^2 = R^2 \tag{1}$$

给出. 在球面坐标 r, θ, φ 下, 这个球面由方程 $r = R, \theta, \varphi$ 任意给出. 所以参数 (θ, φ) 可以作为球面上的局部坐标. 因为要去掉它的北极和南极点 (即 $\theta = 0$ 和 $\theta = \pi$; 这是球面坐标系的奇点; 参看 §1). 已知 (§3.1) 在球面坐标下的欧氏度量 $dx^2 + dy^2 + dz^2$ 取下面的形式:

$$dx^2 + dy^2 + dz^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2). \tag{2}$$

在曲面 $r=R$ 上微分 dr 化为零, 从而对球面在坐标 (θ,φ) 下的度量有形式

$$dl^2=R^2(d\theta^2+\sin^2\theta d\varphi^2). \tag{3}$$

在这里 $0\leqslant\varphi<2\pi, 0\leqslant\theta\leqslant\pi$ (图 11). 在点 0 的小邻域中有 $\sin\theta\approx\theta$. 因此 $\dfrac{dl^2}{R^2}$ 近似地等于欧氏度量 $d\theta^2+\theta^2d\varphi^2$(在极坐标 θ,φ 下). 考虑球面在平面上的球极投影 (参见图 12, 表示了球面的平面截线映射在此平面上). 这里的 (θ,φ) 是球面上的坐标, 而 (r,φ) 是平面上的极坐标. 从图 12 看出有 $\varphi=\varphi, r=R\cot\dfrac{\theta}{2}$. 利用这个变换公式我们可以把度量 (3) 用坐标 r,φ(或是以坐标 x,y) 重写为

$$dl^2=\frac{4R^4}{(R^2+r^2)^2}(dr^2+r^2d\varphi^2)=\frac{4R^4}{(R^2+x^2+y^2)^2}(dx^2+dy^2). \tag{4}$$

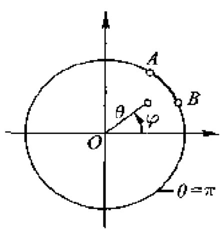


图 11 沿着半径为 $\theta=\pi$ 的圆周所量出的点 A 与 B 之间的长等于 0, 就是说圆盘的边缘粘合成一个点, 这给出了一个二维球面.

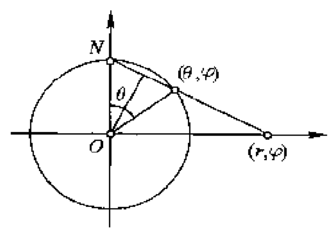


图 12 球极投影:$(\theta,\varphi)\to(r,\varphi)$.

显然, 球面的度量由平面的欧氏度量乘以函数 $\dfrac{4R^4}{(R^2+x^2+y^2)^2}$ 得到, 即

$$dl_{\text{球面}}=\frac{4R^4}{(R^2+x^2+y^2)^2}dl^2_{\text{平面}}. \tag{5}$$

例 9.1. 求在球面上半径为 ρ 的圆盘的周长和面积. 设圆周的中心在北极 N (点 $\theta=0$; 图 13). 而圆周的半径 $\rho=R\theta_0$. 因此圆周的方程为 $\theta=\rho/R$, $0\leqslant\varphi<2\pi$. 半径为 ρ 的圆盘为区域 $\theta\leqslant\rho/R, 0\leqslant\varphi<2\pi$. 在曲线 $\theta=\rho/R=$ 常数上我们有

$$dl^2=R^2(d\theta^2+\sin^2\theta d\varphi^2)=R^2\sin^2\frac{\rho}{R}d\varphi^2,$$

因此长 l_ρ 等于

$$l_\rho=\int_0^{2\pi}R\sin\frac{\rho}{R}d\varphi=2\pi R\sin\frac{\rho}{R}. \tag{6}$$

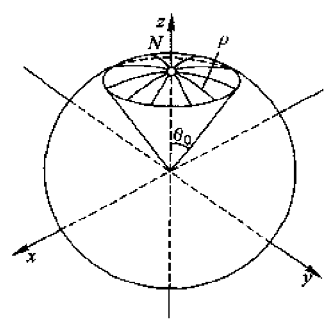


图 13 球面上半径为 ρ 的圆盘

我们看到, 当 $\rho=\pi R/2$ 时长度最大 (赤道); 当 $\rho=\pi R$ 长度为 0 (南极点). 由公式 (6) 得到在球面上圆周长 l_ρ 和半径 ρ 的比值总小于 2π :

$$\frac{l_\rho}{\rho}=2\pi\frac{\sin(\rho/R)}{\rho/R}<2\pi,\quad \rho>0.$$

求半径为 ρ 的圆盘的面积 σ_ρ. 由 dl^2 的表达式 (3) 得到 $\sqrt{g}=R^2|\sin\ \theta|$. 因此面积

$$\begin{aligned}\sigma_\rho&=\iint\limits_{0\leqslant\theta\leqslant\rho/R}R^2|\sin\ \theta|d\theta d\varphi=R^2\int_0^{\rho/R}\sin\ \theta d\theta\int_0^{2\pi}d\varphi\\&=2\pi R^2\left(1-\cos\frac{\rho}{R}\right),\quad \rho\leqslant\frac{\pi}{R}.\end{aligned}\tag{7}$$

当 $\rho=\pi R$ 时我们的圆盘与整个球面吻合, 从而我们得到了球面面积等于 $4\pi R^2$.

如果半径 ρ 小, 则 $\sin\frac{\rho}{R}\approx\frac{\rho}{R}$, 且 $1-\cos\frac{\rho}{R}\approx\frac{\rho^2}{2R^2}$. 我们得到

$$l_\rho=2\pi R\sin\frac{\rho}{R}\approx 2\pi R,\quad \sigma_\rho=2\pi R^2\left(1-\cos\frac{\rho}{R}\right)\approx\pi\rho^2,$$

即在小半径下我们得到此长度和面积近似地有与欧几里得几何中同样的公式.

现在来求半径为 R 的球面的高斯曲率和平均曲率. 我们注意到球面的法截线是半径为 R 的圆 (称其为球面的大圆). 因此任意法截线的曲率为常数, 它等于 R^{-1}(参看 §5.1). 这样, 第二基本型的两个特征值相等, 等于 R^{-1}. 就是说, 球面的高斯曲率等于 $1/R^2$, 而平均曲率等于 $2/R$.

转向球面度量的运动群. 所有绕坐标原点的空间旋转都把半径为 R 的球面变到自己. 这个旋转由正交矩阵给出, 它保持了球面上的黎曼度量不变. 因此球面 S^2 的运动由正交矩阵确定. 回忆 §4.3, 我们记正交矩阵群为 $O(3)$, 称其为完全正交群. 每个旋转可以用三个参数来描述, 因此, 球面的度量至少是运动的三维集合.

重要注解 我们已看到, 球面的所有可能的变换, 当由正交矩阵给出时, 是个运动. 因此群 $O(3)$ 包含了球面 S^2 的所有运动. 但是我们目前还不能证明群 $O(3)$ 实际上就是所有运动的群. 但这个结论是对的. 任何保持 S^2 上度量不变的变换必是 $\mathbb{R}^3$ 中的线性和正交的变换. 然而对此事实的严格证明要求引进测地线的概念. 我们将在第四章中回到这个问题.

§10. 在伪欧氏空间中的类空曲面

1. 伪球面

考虑坐标为 (t,x,y) 的三维伪欧氏空间, 其中的伪欧氏度量为

$$dl^2=dt^2-dx^2-dy^2.\tag{1}$$

半径为 R 的在 $\mathbb{R}_1^3$ 中的伪球面由方程

$$t^2 - x^2 - y^2 = R^2 \tag{2}$$

给出. 这是在三维空间中的双叶双曲面 (图 14). 伪球面整个落在光锥 $t^2-x^2-y^2=0$ 的内部, 从而在伪球面坐标 ρ, χ, φ(见 (3.11),(3.12)) 下由方程

$$\rho = R \text{ (上叶)},$$

$$\rho = -R \text{ (下叶)}$$

给出. 我们将在后面只考虑双曲面的上叶, 即 $\rho = R$. 回忆在伪欧氏坐标 ρ, χ, φ 下度量 $dt^2 - dx^2 - dy^2$ 具有形式 (见 3.12)

$$dl^2 = -\rho^2(d\chi^2 + \mathrm{sh}^2\chi d\varphi^2) + d\rho^2. \tag{3}$$

因此在双曲面上有

$$-dl^2 = R^2(d\chi^2 + \mathrm{sh}^2\chi d\varphi^2), \tag{4}$$

其中 $d\rho = 0$. 这样, 伪欧氏 (即不定的) 度量在伪球面上诱导的度量为负定. 它等价于伪球面 $t^2 - x^2 - y^2 = R^2$ 在 $\mathbb{R}_1^3$ 中是类空的双曲面: 这个曲面的切向量总是类空的 (在其上 $dl^2 < 0$).

定义 10.1. 称度量 (4) 为*罗巴切夫斯基度量*.

以类比于上一节中的方法可以建立伪球面到 (x, y) 平面的球极投影. 伪球面的中心为原点 O, 北极为坐标是 $(R, 0, 0)$ 的点, 南极是坐标为 $(-R, 0, 0)$ 的点. 我们以开圆盘 $x^2 + y^2 < R^2$ 为伪球的上半部 (图 15). 假设 (t, x, y) 为伪球面上点 P (其中 $t > 0$) 的坐标, 并设 (u, v) 为点 $f(P)$ 的坐标, 其中 f 为球极投影. 我们要算出这两种坐标之间的显式联系. 由图 15 可明显看出

$$\frac{x}{u} = \frac{t+R}{R}, \quad \frac{y}{v} = \frac{t+R}{R},$$

由此有 $x = u(1 + t/R), y = v(1 + t/R)$. 将 x 和 y 的这个表达式代入曲面方程 $t^2 - x^2 - y^2 = R^2$ 得到

$$t = -R\left(1 + \frac{2R^2}{u^2 + v^2 - R^2}\right), \tag{5}$$

由此,

$$x = \frac{2R^2 u}{R^2 - u^2 - v^2}, \quad y = \frac{2R^2 v}{R^2 - u^2 - v^2}. \tag{6}$$

公式 (5) 和 (6) 就是我们所要的关于球极投影的公式. 现在可以在坐标 (u, v) 下求伪球面上度量 (4) 的表达式了. 直接的计算表明

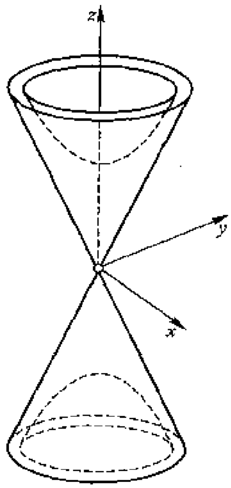


图 14

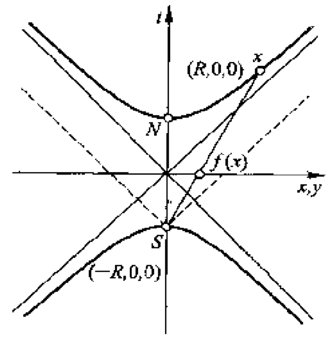


图 15

$$-dl^2 = -(dt^2 - dx^2 - dy^2) = \frac{4R^4}{(u^2+v^2-R^2)^2}(du^2+dv^2). \tag{7}$$

另外 (像对球面 S^2 那样), 在坐标 u,v 下的伪球面的度量与欧几里得平面的度量成比例. 带有负号的在 $u^2+v^2<R^2$ 这个圆盘上的度量 (7) 被称做罗巴切夫斯基几何的庞加莱度量模型. 如果在圆盘 $u^2+v^2<R^2$ 上引进极坐标 (r,φ), 则罗巴切夫斯基度量被写成

$$dl^2 = \frac{4R^4}{(R^2-r^2)^2}(dr^2+r^2d\varphi^2). \tag{8}$$

显然它类似于球面 S^2 的度量 (参看 (9.4)).

我们把球面 S^2 的度量形态与罗巴切夫斯基平面 L^2 的集中放在一张表里 $(R=1)$:

S^2	L^2
$d\theta^2+\sin^2\theta(d\varphi)^2$	$d\chi^2+\mathrm{sh}^2\chi(d\varphi)^2$
$4\dfrac{dx^2+dy^2}{(1+x^2+y^2)^2}$	$4\dfrac{dx^2+dy^2}{(1-x^2-y^2)^2}, x^2+y^2<1$

我们考虑另外一个罗巴切夫斯基度量的表达形式. 由复变函数论中众所周知的事实知道, 存在复平面的分式线性变换把上半平面变为单位圆盘. 我们给出这些变换

中的一个:$z=(1+iw)/(1-iw)$(图 16). 如果我们令 $z=u+iv, w=x+iy$, 于是我们在单位圆盘上引进了新的坐标 $(x,y), y>0$. 直接的计算指出, 罗巴切夫斯基度量在坐标 (x,y) 下有形式

$$dl^2=\frac{dx^2+dy^2}{y^2},\quad y>0. \qquad (9)$$

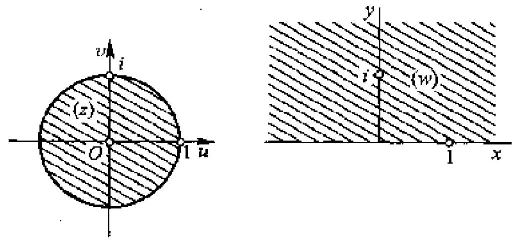


图 16

称度量 (9) 为罗巴切夫斯基几何的克莱因度量模型.

我们来找出罗巴切夫斯基平面的运动群. 空间 $\mathbb{R}_1^3$ 的任意伪正交变换 (见 §6.2) 保持形式 $t^2-x^2-y^2$ 不变, 所以把伪球面 $t^2-x^2-y^2=R^2$ 变到自己. 但是 $O(1,2)$ 中的变换能交换伪球面的上半部和下半部. 因此罗巴切夫斯基平面包含了 $O(1,2)$ 中的正时序变换. 在第四章中将指出这两个群相同. 于是罗巴切夫斯基平面就像球面和欧氏平面一样是具有运动的三参数的集合.

2. $\mathbb{R}_1^3$ 中类空曲面的曲率

$\mathbb{R}_1^3$ 中一个曲面是类空的是说它的每个切向量是类空的. 换句话说, 闵可夫斯基度量 $dl^2=dt^2-dx^2-dy^2$ 在曲面上诱导出负定的度量.

对于曲面由函数 $t=f(x,y)$ 的图像给出, 在曲面上的度量有形式

$$-dl^2=-(dt^2-dx^2-dy^2)=(1-f_x^2)dx^2-2f_xf_ydxdy+(1-f_y^2)dy^2=g_{ij}dx^idx^j,$$
$$x^1=x,\quad x^2=y.$$

我们有 $\det(g_{ij})=1-f_x^2-f_y^2$; 曲面的单位法向量 $m=\dfrac{(1,f_x,f_y)}{\sqrt{1-f_x^2-f_y^2}}$(类时的). 曲面的第二基本型由等式

$$b_{ij}=\langle\frac{\partial^2 r}{\partial x^i\partial x^j},m\rangle,\quad i,j=1,2$$

定义. 令

$$K=-\frac{\det\, b_{ij}}{\det\, g_{ij}} \qquad (10)$$

来引进类空曲面的高斯曲率. 那么对于由方程 $t=f(x,y), 1-f_x^2-f_y^2>0$ 的曲面, 则由类似于 §8 那样, 有

$$K=\frac{f_{xy}^2-f_{xx}f_{yy}}{(1-f_x^2-f_y^2)^2}. \qquad (11)$$

特别地, 对于双曲面 $t^2-x^2-y^2=1$(罗巴切夫斯基几何) 我们得到了 $K\equiv-1$.

注 1. 在第四章中我们将从内蕴几何出发计算罗巴切夫斯基平面的曲率. 我们重新得到 $K\equiv-1$; 并解释了定义 (10) 中符号选取的原因.

2. 像在 §8 中那样, 我们得到: 高斯曲率 $K<0$ 的类空曲面在 $\mathbb{R}_1^3$ 中是凸的.

§11. 几何中的复语言

1. 复坐标和实坐标

在许多几何问题中常常使用复数的语言表示. 为此我们要在这里解释一些我们所需要的简单事实.

假设给出了以 $e_1, \cdots, e_n$ 为基的复数域 $\mathbb{C}$ 上的 n 维线性空间 $\mathbb{C}^n$. 任意向量 $\xi \in \mathbb{C}^n$ 有形式

$$\xi = z^k e_k, \quad z^k = x^k + iy^k, \tag{1}$$

其中 z^k 为复坐标. 空间 $\mathbb{C}^n$ 可以看作为 $2n$ 维的线性空间 $\mathbb{R}^{2n}$, 这是实数域上的空间, 其中在 $\mathbb{R}^{2n}$ 中的基由

$$e_1, \cdots, e_n, ie_1, \cdots, ie_n \tag{2}$$

给出. 于是我们有

$$\xi = z^k e_k = x^k e_k + y^k(ie_k), \tag{3}$$

其中 $(x^1, \cdots, x^n, y^1, \cdots, y^n)$ 为向量 ξ 的实坐标. 所描述的这个运作称为实 (数) 化.

复空间 $\mathbb{C}^n$ 的复线性非异变换构成群 $GL(n, \mathbb{C})$. 这是 $n \times n$ 阶其行列式不为零的复矩阵的群. 在实数化下, 每个这种变换给出了实空间 $\mathbb{R}^{2n}$ 中的一个线性变换. 我们因而得到了一个实化映射

$$r : GL(n, \mathbb{C}) \to GL(2n, \mathbb{R}). \tag{4}$$

例 11.1. 设 $n = 1$. 我们有一维的复空间, 坐标 $z = x + iy$. 空间 $\mathbb{C}$ 的线性变换为乘以复数 $\lambda \neq 0$:

$$z \mapsto \lambda z. \tag{5}$$

如果 $\lambda = a + ib, a^2 + b^2 \neq 0$, 则得到

$$z = x + iy \mapsto (a + ib)(x + iy) = (ax - by) + i(bx + ay).$$

因此对应的空间 $\mathbb{R}^2$ 的变换由矩阵

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = r(\lambda), \quad a^2 + b^2 \neq 0 \tag{6}$$

给出. 显然, 由空间 $\mathbb{C}$ 的复线性变换得到的不是空间 $\mathbb{R}^2$ 的任意的线性变换.

类比地不难证明, 如果在 n 维的情形, A 为 $GL(n\mathbb{C})$ 中矩阵, 那么 $A = A + iB$, 其中 A 和 B 为实矩阵, 于是在映射 $r : GL(n, \mathbb{C}) \to GL(2n, \mathbb{R})$ 下我们有

$$r(A) = \begin{pmatrix} A & B \\ -B & A \end{pmatrix}. \tag{7}$$

习题 11.1. 证明矩阵 $r(A)$ 的行列式为 $\det(r(A)) = |\det A|^2$.

注 我们要指出映射 $r : GL(n,\mathbb{C}) \to GL(2n,\mathbb{R})$ 的另一种描述. 如果 $A \in GL(n,\mathbb{C})$ 为复线性变换, 则对任意向量 ξ 有

$$A(i\xi) = iA(\xi). \tag{8}$$

在实化下, 乘以 i 可由 $2n$ 维实空间中线性变换 I 来实现, 即

$$r(i) = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \quad 其中 I = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix}. \tag{9}$$

因为

$$i(e_k) = ie_k, \quad i(ie_k) = -e_k,$$

故在基 (2) 下乘以 i 的运算 $I = r(i)$ 的矩阵有 (9) 的形式. 由 (8) 知道, 矩阵 I 与矩阵 $r(A)$ 可交换. 这表明此运算的交换性.

在 $GL(n,\mathbb{C})$ 中有行列式为 1 的矩阵组成的子群, 记其为 $SL(n,\mathbb{C})$.

2. 埃尔米特内积

在复数情形下我们给出下面的空间 $\mathbb{C}^n$ 中的内积:

$$\langle \xi_1, \xi_2 \rangle_{\mathbb{C}} = \sum_{k=1}^{n} z_1^k \bar{z}_2^k, \quad \langle \xi, \xi \rangle_{\mathbb{C}} = \sum_{k=1}^{n} |z^k|^2; \tag{10}$$

这里 $\xi_1 = z_1^i e_i$, 等等, 而 "$-$" 表示复共轭. 它有下列性质:

$$\begin{gathered} \langle \lambda\xi, \eta \rangle_{\mathbb{C}} = \lambda \langle \xi, \eta \rangle_{\mathbb{C}}, \\ \langle \xi, \lambda\eta \rangle_{\mathbb{C}} = \bar{\lambda} \langle \xi, \eta \rangle_{\mathbb{C}}, \\ \langle \xi, \eta \rangle_{\mathbb{C}} = \overline{\langle \eta, \xi \rangle}_{\mathbb{C}}, \\ \langle \xi_1 + \xi_2, \eta \rangle_{\mathbb{C}} = \langle \xi_1, \eta \rangle_{\mathbb{C}} + \langle \xi_2, \eta \rangle_{\mathbb{C}}, \\ \langle \xi, \xi \rangle_{\mathbb{C}} > 0, \quad 其中 \xi \neq 0. \end{gathered} \tag{11}$$

任意具性质 (11) 的内积被称为埃尔米特内积.

在实化空间 $\mathbb{R}^{2n}$ 中也可以引进欧氏内积, 即如果 $\xi_1 = (x_1^1, \cdots, x_1^n, y_1^1, \ldots, y_1^n)$, $\xi_2 = (x_2^1, \cdots, x_2^n, y_2^1, \cdots, y_2^n)$, 则令

$$\langle \xi_1, \xi_2 \rangle_{\mathbb{R}} = \sum_{k=1}^{n} (x_1^k x_2^k + y_1^k y_2^k). \tag{12}$$

我们指出, 上面的埃尔米特内积与这个欧氏内积的关联:

$$\mathrm{Re}\langle\xi_1,\xi_2\rangle_{\mathbb{C}} = \langle\xi_1,\xi_2\rangle_{\mathbb{R}}, \tag{13}$$

其中 Re 表示复数的实部. 事实上, 我们有

$$\begin{aligned}\mathrm{Re}\langle\xi_1,\xi_2\rangle_{\mathbb{C}} &= \mathrm{Re}\sum_{k=1}^{n}(x_1^k+iy_1^k)(x_2^k-iy_2^k)\\ &= \sum_{k=1}^{n}(x_1^kx_2^k+y_1^ky_2^k).\end{aligned}$$

特别, 由公式 (10) 得出: 向量自身的埃尔米特内积同于其欧几里得内积:

$$\langle\xi,\xi\rangle_{\mathbb{C}} = \langle\xi,\xi\rangle_{\mathbb{R}}. \tag{14}$$

设 $A\in GL(n,\mathbb{C})$ 为复非异线性变换.

定义 11.1. 如果对任意向量 ξ_1,ξ_2 有

$$\langle A\xi_1,A\xi_2\rangle_{\mathbb{C}} = \langle\xi_1,\xi_2\rangle_{\mathbb{C}}, \tag{15}$$

则称 A 为*酉变换*.

如果在基 $e_1,\cdots,e_n$ 下, 埃尔米特内积有形式 (10), 而 A 由矩阵 $A=(\lambda_k^l)$ 给出. 于是由酉变换的条件得到

$$\sum_{k=1}^{n}\lambda_k^l\bar{\lambda}_k^{l'} = \delta^{ll'}, \tag{16}$$

或者写成矩阵形式:

$$A\overline{A}^{\mathrm{T}} = 1 \longleftrightarrow \overline{A}^{\mathrm{T}} = A^{-1} \tag{17}$$

("T" 表示转置). 酉矩阵构成一个群, 以 $U(n)$ 表示, 称之为*酉群*. 由 (17) 得出

$$\det(\overline{A}^{\mathrm{T}}A) = (\overline{\det A})(\det A) = |\det A|^2 = 1.$$

因此, 酉矩阵的行列式的绝对值等于 1. 群 $U(n)$ 中有子群 $SU(n)$, 它由行列式为 1 的酉矩阵组成.

注 可以看出, 类似于 §4 那样, $U(n)$ 是 $\mathbb{C}^n$ 中酉度量下的运动群.

考虑在 $GL(2n,\mathbb{R})$ 中的像 $r(U(n))$. 由公式 (13) 知道, 如果 $A\in U(n)$, 即 A 保持了自身埃尔米特内积不变, 则矩阵 $r(A)$ 也保持了 $\mathbb{R}^{2n}$ 中的自身的欧几里得内积不变. 因而酉群在 $GL(2n,\mathbb{R})$ 中的像 $r(U(n))$ 有形式 (也可见习题 11. 1)

$$r(U(n)) = SO(2n)\bigcap r(GL(n,\mathbb{C})). \tag{18}$$

我们也可以用伪欧氏空间来类比地研究伪埃尔米特空间 $\mathbb{C}^n_{p,q}, p+q=n$, 其中具坐标 $(z^1,\cdots,z^n)$ 的向量 ξ 的长度平方由公式

$$\langle\xi,\xi\rangle_{p,q}=|z^1|^2+\cdots+|z^p|^2-\cdots-|z^n|^2 \tag{19}$$

给出. 我们用 $U(p,q)$ 表示使形式 (19) 保持不变的复线性变换, 而用 $SU(p,q)$ 表示矩阵行列式为 1 的 $U(p,q)$ 的子群.

3. 复变换群的例子

我们在前面看到群 $GL(1,\mathbb{C})$ 是非零复数 (在乘法下) 的群. 群 $U(1)$ 由所有模为 1 的复数组成: $U(1)=\{e^{i\varphi}\}$. 我们注意到有 $r(e^{i\varphi})=\begin{pmatrix}\cos\varphi & \sin\varphi\\ -\sin\varphi & \cos\varphi\end{pmatrix}$, 因此映射 r 给定了群 $U(1)$ 和 $SO(2)$ 的同构.

现在来考虑群 $SL(2,\mathbb{C})$. 设 $A=\begin{pmatrix}a & b\\ c & d\end{pmatrix}\in SL(2,\mathbb{C})$, 即 $ad-bc=1$. 作 (扩充了的) 复平面 $\mathbb{C}$ 上对应于矩阵 A 的分式线性变换:

$$z'=\frac{az+b}{cz+d}. \tag{20}$$

如果 $A'=\begin{pmatrix}a' & b'\\ c' & d'\end{pmatrix}$ 为另一个行列式为 1 的矩阵, 则有

$$z''=\frac{a'z'+b'}{c'z'+d'}=\frac{(a'a+b'c)z+a'b+b'd}{(c'a+d'c)z+c'b+d'd},$$

即所构造的映射代表了一个同态

$$\varphi: SL(2,\mathbb{C})\mapsto L, \tag{21}$$

这里的 L 表示分式线性变换群. 不难看出, 同态 φ 的核由两个矩阵构成: $\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$ 和 $\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$, 同时 φ 把 $SL(2,\mathbb{C})$ 映满了整个群 L (满射). 所以

$$L=SL(2,\mathbb{C})/\pm 1. \tag{22}$$

再考虑群 $U(2)$. 如果矩阵 $\begin{pmatrix}a & b\\ c & d\end{pmatrix}$ 属于 $U(2)$, 则

$$|a|^2+|b|^2=1,\quad |c|^2+|d|^2=1,\quad a\bar{c}+b\bar{d}=0. \tag{23}$$

它的子群 $SU(2)$ 由补充的条件 $ad-bc=1$ 从其中得到. 因此 $SU(2)$ 由下面形式的矩阵构成:

$$\begin{pmatrix}a & b\\ -\bar{b} & \bar{a}\end{pmatrix},\quad |a|^2+|b|^2=1. \tag{24}$$

另一个例子是群 $SU(1,1)$. 它由形如

$$\begin{pmatrix} c & d \\ \bar{d} & \bar{c} \end{pmatrix}, \quad |c|^2 - |d|^2 = 1 \tag{25}$$

的矩阵组成. 注意, 其中 $c \neq 0$(因为 $|c| \geqslant 1$).

存在由下面公式定义的映射 $SU(1,1) \to SU(2)$:

$$\begin{pmatrix} c & d \\ \bar{d} & \bar{c} \end{pmatrix} \mapsto \begin{pmatrix} 1/c & d/c \\ -\bar{d}/\bar{c} & 1/\bar{c} \end{pmatrix} = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}. \tag{26}$$

逆映射在 $a \neq 0$ 时有定义. 这个映射并不是个群同态.

§12. 解析函数

1. 长度元和函数微分的复表示

设给出了在复坐标下的空间 $\mathbb{C}^n$ 中的曲线, 其形式为

$$z^k = z^k(t) = x^k(t) + iy^k(t). \tag{1}$$

这时在实坐标 (x^k, y^k) 下, 我们得到曲线 $x^1(t), \cdots, x^n(t), y^1(t), \cdots, y^n(t)$. 它的长度有形式

$$l = \int_{t_0}^{t_1} \sqrt{\sum_{k=1}^{n} (\dot{x}^k)^2 + (\dot{y}^k)^2} dt = \int_{t_0}^{t_1} \sqrt{\sum_{k=1}^{n} \dot{z}^k \dot{\bar{z}}^k} dt. \tag{2}$$

在空间 $\mathbb{R}^{2n}$ 中通过下面的公式容易从坐标 (x^k, y^k) 过渡到复坐标 $z^k, \bar{z}^k, k = 1, \cdots, n$:

$$\begin{aligned} z^k &= x^k + iy^k, \quad \bar{z}^k = x^k - iy^k \\ x^k &= \frac{1}{2}(z^k + \bar{z}^k), \quad y^k = \frac{1}{2i}(z^k - \bar{z}^k). \end{aligned} \tag{3}$$

这时长度元的复形式可写成

$$dl^2 = \sum_{k=1}^{n} dz^k d\bar{z}^k, \tag{4}$$

其中令

$$dz^k = dx^k + idy^k, \quad d\bar{z}^k = dx^k - idy^k. \tag{5}$$

我们在 $\mathbb{C}^n$ 上的复函数空间中引进算子

$$\begin{aligned} \frac{\partial}{\partial z^k} &= \frac{1}{2}\left(\frac{\partial}{\partial x^k} - i\frac{\partial}{\partial y^k}\right), \\ \frac{\partial}{\partial \bar{z}^k} &= \frac{1}{2}\left(\frac{\partial}{\partial x^k} + i\frac{\partial}{\partial y^k}\right). \end{aligned} \tag{6}$$

显然, 我们有

$$\begin{aligned}\frac{\partial}{\partial x^k} &= \frac{\partial}{\partial z^k} + \frac{\partial}{\partial \bar{z}^k},\\ \frac{\partial}{\partial y^k} &= i\left(\frac{\partial}{\partial z^k} - \frac{\partial}{\partial \bar{z}^k}\right).\end{aligned} \tag{7}$$

注意成立有下面的等式:

$$\begin{aligned}\frac{\partial}{\partial \bar{z}^k}(z^k) &= \frac{\partial}{\partial z^k}(\bar{z}^k) = 0\\ \frac{\partial}{\partial z^k}(z^k) &= \frac{\partial}{\partial \bar{z}^k}(\bar{z}^k) = 1.\end{aligned} \tag{8}$$

由公式 (5) 和 (7) 可以得出下面的论断.

引理 12.1. 任意复值函数 $f(x^1,\cdots,x^n,y^1,\cdots,y^n)$ 的微分有

$$df = \frac{\partial f}{\partial z^1}dz^1 + \cdots + \frac{\partial f}{\partial z^n}dz^n + \frac{\partial f}{\partial \bar{z}^1}d\bar{z}^1 + \cdots + \frac{\partial f}{\partial \bar{z}^n}d\bar{z}^n. \tag{9}$$

现在考虑具有复系数的变量为 $x^1,\cdots,x^n,y^1,\cdots,y^n$ 的任意多项式 $P(x^1,\cdots,x^n,y^1,\cdots,y^n)$. 使用变量替换 (3) 我们从多项式 $P(x^1,\cdots,x^n,y^1,\cdots,y^n)$ 得到了多项式 $Q(z^1,\cdots,z^n,\bar{z}^1,\cdots,\bar{z}^n)$. 成立下面的论断:

定理 12.1. 多项式 $Q(z,\bar{z}) = P(x,y)$ 只依赖于变量 $z^1,\cdots,z^n$, 而与变量 $\bar{z}^1,\cdots,\bar{z}^n$ 无关当且仅当它满足等式

$$\frac{\partial P}{\partial \bar{z}^k} \equiv 0, \quad k = 1,\cdots,n. \tag{10}$$

证明 算子 $\frac{\partial}{\partial z^k}$ 和 $\frac{\partial}{\partial \bar{z}^k}$ 具有下面明显的性质 (莱布尼茨公式):

$$\frac{\partial}{\partial z^k}(fg) = \frac{\partial f}{\partial z^k}g + f\frac{\partial g}{\partial z^k}, \quad \frac{\partial}{\partial \bar{z}^k}(fg) = \frac{\partial f}{\partial \bar{z}^k}g + f\frac{\partial g}{\partial \bar{z}^k}. \tag{11}$$

另外, 利用公式 (8), 由其得到

$$\frac{\partial}{\partial \bar{z}^k}[(z^k)^m] = 0, \quad \frac{\partial}{\partial \bar{z}^k}[(\bar{z}^k)^m] = m(\bar{z}^k)^{m-1}. \tag{12}$$

于是立即知道, 如果 $P(x,y) = Q(z,\bar{z})$ 不依赖于 $\bar{z}^k$, 则 $\frac{\partial P}{\partial \bar{z}^k} \equiv 0$.

现在证明相反的断言. 设多项式 P 依赖于 $\bar{z}^k$, 其中 $\bar{z}^k$ 出现在 P 中的最高阶数设为 m. 我们将证明 $\frac{\partial P}{\partial \bar{z}^k} \neq 0$.

多项式 P 具有形式

$$P = A_0(\bar{z}^k)^m + A_1(\bar{z}^k)^{m-1} + \cdots + A_m,$$

其中 $A_0,\cdots,A_m$ 为所有变量 $z^1,\cdots,z^n$ 和除去 $\bar{z}^k$ 以外的所有变量 $\bar{z}^q$ 的多项式. 于

是, $\frac{\partial A_i}{\partial \bar{z}^k} \equiv 0, i=0,\cdots,m$($A_i$ 不依赖于 $\bar{z}^k$). 因此

$$\frac{\partial}{\partial \bar{z}^k} P = A_0 m(\bar{z}^k)^{m-1} + A_1(m-1)(\bar{z}^k)^{m-2} + \cdots.$$

因为 $A_0 \neq 0$, 故 $\frac{\partial P}{\partial \bar{z}^k} \neq 0$. 定理证完. □

注　此定理不仅用于多项式, 而且可用于收敛的幂级数: 它与 $\bar{z}^k$ 无关等价于条件 $\frac{\partial f}{\partial \bar{z}^k} \equiv 0$.

定义 12.1.　称满足等式

$$\frac{\partial f}{\partial \bar{z}^k} \equiv 0, \quad k=1,\cdots,n \tag{13}$$

的函数 $f(x^1,y^1,\cdots,x^n,y^n)$ 为复解析函数.

对于两个实变量的函数 $f(x,y)=f(x(z,\bar{z}),y(z,\bar{z})), z=x+iy, \bar{z}=x-iy$ 而言, 解析条件 (即不依赖于 $\bar{z}$) 具有形式

$$2\frac{\partial f}{\partial \bar{z}} = \frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y} \equiv 0. \tag{14}$$

如果 $f(x,y)=u(x,y)+iv(x,y)$, 则 (14) 可以写为

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \tag{15}$$

称等式 (15) 为柯西 – 黎曼方程. 显然, 由 (15) 得到

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0. \tag{16}$$

因此, 复解析函数的实部和虚部都是拉普拉斯方程的解 (即调和函数): $\Delta u = 0, \Delta v = 0$, 其中 $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = 4\frac{\partial^2}{\partial z \partial \bar{z}}$ 为拉普拉斯算子.

2. 复坐标变换

设在 n 维复空间 $\mathbb{C}^n$ 中的某区域上给出了两组复坐标

$$\begin{aligned} &z^1 = x^1 + iy^1, \quad \cdots, \quad z^n = x^n + iy^n, \\ &w^1 = u^1 + iv^1, \quad \cdots, \quad w^n = u^n + iv^n. \end{aligned} \tag{17}$$

这时坐标 $w^k = u^k + iv^k$ 以坐标 $z^k = x^k + iy^k$ 的函数形式给出:

$$w^k = w^k(x^1,y^1,\cdots,x^n,y^n), \quad k=1,\cdots,n. \tag{18}$$

定义 12.2. 如果

$$\frac{\partial w^k}{\partial \bar{z}^l} \equiv 0, \quad k, l = 1, \cdots, n, \tag{19}$$

则称坐标变换 (18) 为复解析坐标变换.

令

$$a_l^k = \frac{\partial w^k}{\partial z^l}, \quad k, l = 1, \cdots, n, \tag{20}$$

则可引进复坐标变换 (18) 的雅可比矩阵 (a_l^k). 称矩阵 (a_l^k) 的行列式为变换 (18) 的复雅可比:

$$J_{\mathbb{C}} = \det(a_l^k). \tag{21}$$

变换 (18) 在实化空间 $\mathbb{R}^{2n}$ 中给出了变换

$$\begin{aligned} &u^k = u^k(x^1, y^1, \cdots, x^n, y^n), \quad v^k = v^k(x^1, y^1, \cdots, x^n, y^n), \\ &k = 1, \cdots, n. \end{aligned} \tag{22}$$

设

$$J_{\mathbb{R}} = \frac{\partial(u, v)}{\partial(x, y)} \tag{23}$$

为变换 (22) 的 (实) 雅可比.

原来发现复和实的雅可比间关系极其简单.

引理 12.2. 对于复解析坐标变换有下面的等式:

$$J_{\mathbb{R}} = |J_{\mathbb{C}}|^2.$$

证明 设 $A = (a_l^k) = \left(\dfrac{\partial w^k}{\partial z^l}\right)$ 为 (复) 雅可比矩阵, $J_{\mathbb{C}} = \det A$. 我们要求由坐标 $z^1, \cdots, z^n, \bar{z}^1, \cdots, \bar{z}^n$ 到坐标 $w^1, \cdots, w^n, \bar{w}^1, \cdots, \bar{w}^n$ 在空间 $\mathbb{R}^{2n}$ 中变换的实雅可比矩阵. 由复解析条件我们有

$$\frac{\partial w^k}{\partial z^l} = a_l^k, \quad \frac{\partial w^k}{\partial \bar{z}^l} = \frac{\partial \bar{w}^k}{\partial z^l} = 0, \quad \frac{\partial \bar{w}^k}{\partial \bar{z}^l} = \bar{a}_l^k.$$

因此所求的雅可比矩阵具有形式 $\begin{pmatrix} A & 0 \\ 0 & \bar{A} \end{pmatrix}$. 它的行列式等于

$$\det \begin{pmatrix} A & 0 \\ 0 & \bar{A} \end{pmatrix} = |\det A|^2 = |J_{\mathbb{C}}|^2.$$

现在留意一下, 从坐标 $x^1,\cdots,x^n,y^1,\cdots,y^n$ 到坐标 $z^1,\cdots,z^n,\bar{z}_1,\cdots,\bar{z}_n$ 的转移或从 (u,v) 到 $(w,\bar{w})$ 的转移由矩阵

$$B=\left(\begin{array}{ccc|ccc}1 & & 0 & i & & 0\\ & \ddots & & & \ddots & \\ 0 & & 1 & 0 & & i\\ \hline 1 & & 0 & -i & & 0\\ & \ddots & & & \ddots & \\ 0 & & 1 & 0 & & -i\end{array}\right),\quad \det B=(-2i)^n$$

给出. 因而

$$J_{\mathbb{R}}=\det\left(B^{-1}\begin{pmatrix}A & 0\\ 0 & \bar{A}\end{pmatrix}B\right)=|J_{\mathbb{C}}|^2.$$ □

推论 如果变换 (18) 的 (复) 雅可比不为零, 则局部可通过 w 表出 z:

$$z^k=z^k(w^1,\cdots,w^n),\quad k=1,\cdots,n,$$

这时反函数 $z^k(w)$ 是复解析的.

证明 我们已知从 $x^1,\cdots,x^n,y^1,\cdots,y^n$ 到 $u^1,\cdots,u^n,v^1,\cdots,v^n$ 的坐标变换的雅可比矩阵有形式 $B^{-1}\begin{pmatrix}A & 0\\ 0 & \bar{A}\end{pmatrix}B$ (见前面的证明). 因为由所给条件知此矩阵的行列式不为零, 故由逆映射存在定理 (参看 §1) 知局部地存在雅可比矩阵为 $B^{-1}\begin{pmatrix}A^{-1} & 0\\ 0 & \bar{A}^{-1}\end{pmatrix}B$ 的变换. 由此直接知道函数 $z^k(w,\bar{w})$ 存在. 对于在 $\mathbb{R}^{2n}$ 的变换 $(w,\bar{w})\to(z,\bar{z})$ 有形式

$$\begin{pmatrix}A^{-1} & 0\\ 0 & \bar{A}^{-1}\end{pmatrix}.$$

引理得证. □

例 12.1. 当 $n=1$ 时变换 (18) 有形式

$$w=w(z),\quad \frac{\partial w}{\partial\bar{z}}=0. \tag{24}$$

这个变换在 $\dfrac{\partial w}{\partial z}\neq 0$ 处, 可逆.

在复的情形中, 变换总意味着从空间 $\mathbb{C}^n$ 的一个区域到另一个区域的相互一一的映射, 而这个映射由复解析函数给出.

$n=1$ 时的变换的例子.

例 12.2. 复仿射变换

$$w(z)=az+b,\quad a\neq 0. \tag{25}$$

在这个变换下 $\dfrac{dw}{dz}=a\neq 0.$

回忆 §4.2, 以实的观点看,(25) 给出了平面的运动同时保持了定向的伸缩.

例 12.3. 分式线性变换

$$w(z)=\frac{az+b}{cz+d},\quad ad-bc\neq 0 \tag{26}$$

(可以假定 $ad-bc=1$). 这时

$$\frac{dw}{dz}=-\frac{1}{(cz+d)^2}\neq 0. \tag{27}$$

严格地说, 变换 (26) 在 $z=-\dfrac{d}{c}$ 没有定义. 可以认为 (目前完全是形式地) 变换 (26) 是在扩张的复平面上给出的, 在那里复直线 $\mathbb{C}$ 上附加了一个无穷远点 ∞, 这时

$$-\frac{d}{c}\mapsto\infty,\quad \infty\mapsto\frac{a}{c}. \tag{28}$$

例如, 变换

$$w=\frac{1+iz}{1-iz} \tag{29}$$

把上半平面 $\operatorname{Im} z>0$ 变到了单位圆 $|w|<1$. 我们曾在 §10 利用过此变换来构造罗巴切夫斯基几何的克莱因模型.

3. 复空间中的曲面

我们考虑在二维复空间 $\mathbb{C}^2$ 中的一维曲面 (复曲线) 这个最简单的情形. 这样的曲线在空间 $\mathbb{C}^2$ 中的坐标 (w,z) 下由方程

$$f(w,z)=0 \tag{30}$$

给出, 其中 $f(w,z)$ 为变量 w,z 的复解析函数. 方程 (30) 是有两个实方程的方程组:$u=0,v=0$, 其中 $f=u+iv$, 因此给出了 $\mathbb{C}^2=\mathbb{R}^4$ 中的一个二维曲面. 令

$$\operatorname{grad}_{\mathbb{C}}f=\left(\frac{\partial f}{\partial w},\frac{\partial f}{\partial z}\right), \tag{31}$$

由此引进复梯度 $\operatorname{grad}_{\mathbb{C}} f$.

如果在点 (w_0,z_0) 上有 $\operatorname{grad}_{\mathbb{C}}f|_{w_0,z_0}\neq 0$, 则称曲线 (30) 上的点 (w_0,z_0) 为非异点. 我们有下面的类比于隐函数定理的复形式 (只叙述, 不证明).

定理 12.2. 设 $f(w,z)$ 是变量 w,z 的复解析函数, 并且 $\operatorname{grad}_{\mathbb{C}} f$ 在使 $f(w_0,z_0)=0$ 的点 (w_0,z_0) 上不为 0.

假定，例如是 $\dfrac{\partial f}{\partial w}$ 在此点不为零. 于是在点 (w_0,z_0) 的充分小的邻域内方程 $f(w,z)=0$ 有唯一的复解析的解 $w=w(z)$，使得 $f(w(z),z)=0, w_0=w(z_0), \dfrac{\partial w}{\partial\bar z}\equiv 0$.

例 12.4. 设 $f(w,z)$ 为两个变量的多项式. 称方程 $f(w,z)=0$ 的形如 $w=w(z)$ 的全部解的集合为*多值代数函数*，而称曲面 (复曲线)$f(w,z)=0$ 为这个多值函数的*图像*或*黎曼面*①.

重要的特殊情形是*超椭圆曲线*，即由方程

$$f(w,z)=w^2-P_n(z)=0 \tag{32}$$

给出的黎曼面，其中 $P_n(z)$ 为 n 次多项式. 这是代数函数 $w=\sqrt{P_n(z)}$ 的图像.

引理 12.3. 曲面 (32) 非异当且仅当 $P_n(z)$ 没有重根.

证明 计算函数 f 的梯度：

$$\mathrm{grad}_{\mathbb{C}}f=\left(\frac{\partial f}{\partial w},\frac{\partial f}{\partial z}\right)=\left(2w,-\frac{dP_n(z)}{dz}\right).$$

如果 $\mathrm{grad}_{\mathbb{C}}f$ 在曲面 (32) 上有零点，则

$$2w=0,\quad \frac{dP_n(z)}{dz}=0,\quad w^2-P_n(z)=0.$$

由此得出奇点的坐标为 $(0,z_0)$，而 z_0 为多项式 $P_n(z)$ 和它的导数 $\dfrac{dP_n}{dz}$ 的公共根. 没有公共根等价于多项式 $P_n(z)$ 没有重根. 引理证完. □

在曲面 (32) 上，根据复的隐函数定理可以在 $\dfrac{\partial f}{\partial w}=2w\neq 0$，即 $P_n(z)\neq 0$ 的点上引进局部坐标 z. 如果 $P_n(z)=0$，则 $\dfrac{\partial f}{\partial z}=-\dfrac{dP_n}{dz}\neq 0$. 那么在这点可取 w 为局部坐标.

再回到任意非异复解析曲线 $f(w,z)=0$ 的情形. 设 $\left.\dfrac{\partial f}{\partial w}\right|_{(w_0,z_0)}\neq 0$ 于点 (w_0,z_0). 于是 $w=w(z), \dfrac{\partial w}{\partial\bar z}\equiv 0$. 在 $\mathbb{C}^2$ 中的度量元 dl^2 为

$$dl^2=dwd\bar w+dzd\bar z, \tag{33}$$

它在曲面 $f(w,z)=0$ 转化为

$$dl^2=dwd\bar w+dzd\bar z=\left(1+\left|\frac{dw}{dz}\right|^2\right)dzd\bar z. \tag{34}$$

①这是通常使用的黎曼面定义 (譬如参看 [16]) 的一个简化形式. 在曲面 $f(w,z)=0$ 无奇点和无自交的情形，我们的定义等价于一般的情形.

如果 $z=x+iy$, 则在曲面 $f(w,z)=0$ 上长度元的平方有形式

$$dl^2 = h(z,\bar{z})dzd\bar{z} = g(x,y)(dx^2+dy^2), \tag{35}$$

其中 $h(z,\bar{z}) = g(x,y) = 1 + \left|\dfrac{dw}{dz}\right|^2$.

定义 12.3. 如果曲面上的度量在坐标 x,y 下具有 $dl^2 = g(x,y)(dx^2+dy^2)$ 的形式, 我们称此坐标为共形坐标.

有简单的引理:

引理 12.4. 度量的共形形式 (只) 对于复解析的坐标变换和它们与复共轭的复合是不变的.

证明 假设对坐标 z 度量有形式

$$dl^2 = h(z,z)dzd\bar{z}.$$

设 $z = z(w)$, $\dfrac{\partial z}{\partial \bar{w}} \equiv 0$. 于是

$$dz = \left(\frac{dz}{dw}\right)dw, \quad d\bar{z} = \overline{\left(\frac{dz}{dw}\right)}d\bar{w}$$

和

$$dl^2 - h(z,\bar{z})dzd\bar{z} = h(z(w,\bar{w}),\overline{z(w,\bar{w})})\left|\frac{dz}{dw}\right|^2 dwd\bar{w}.$$

如果 $\dfrac{\partial z}{\partial w}=0$, 证明完全类似. 如果 $z=z(w,w)$ 而 $\dfrac{\partial z}{\partial w}\neq 0$, $\dfrac{\partial z}{\partial \bar{w}} \neq 0$, 则容易看出, 在变量 $w,\bar{w}$ 下 dl^2 有形式

$$\begin{aligned} dl^2 &= hdzd\bar{z} = h\left|\frac{\partial z}{\partial w}dw + \frac{\partial z}{\partial \bar{w}}d\bar{w}\right|^2 \\ &= h(a_{11}(dx')^2 + 2a_{12}dx'dy' + a_{22}(dy')^2), \end{aligned}$$

其中 $w = x'+iy'$; 显然 $a_{12}\neq 0$. 引理得证. □

§13. 曲面度量的共形形式

1. 等温坐标、共形坐标下的高斯曲率

设在 $\mathbb{R}^3$ 中给出由参数定义的二维曲面

$$x = x(p,q), \quad y = y(p,q), \quad z = z(p,q), \tag{1}$$

其中 p,q 在空间 $\mathbb{R}^2$ 中某区域变动. 于是在曲面上存在诱导度量

$$dx^2+dy^2+dz^2 = E(dp)^2 + 2Fdpdq + G(dq)^2, \tag{2}$$

$g = EG - F^2 > 0$. 通过对曲面上局部坐标 (p, q) 的变换, 度量 dl^2 可以化为共形形式. 就是说我们有 (参看 [1]).

定理 13.1. 设 E, F, G 是变量 p, q 的 (实) 解析函数. 于是可以引进新的局部坐标 u, v 使得在此新坐标下度量 dl^2 有下面形式

$$dl^2 = f(u, v)(du^2 + dv^2). \tag{3}$$

注　称这样的坐标为等温坐标或共形坐标. 因此, 在等温坐标下度量有上节所说的意义下的共形的形式.

证明　我们把二次型 $dl^2 = g_{11}dp^2 + 2g_{12}dpdq + g_{22}dq^2$ 分解因子为

$$dl^2 = \left(\sqrt{E}dp + \frac{F + i\sqrt{g}}{\sqrt{E}}dq\right)\left(\sqrt{E}dp + \frac{F - i\sqrt{g}}{\sqrt{E}}dq\right),$$

$$g_{11} = E, \quad g_{12} = F, \quad g_{22} = G, \quad g = EG - F^2.$$

我们要寻找新坐标 u, v, 它们是 p, q 的函数:$u = u(p, q), v = v(p, q)$, 使得度量 (2) 具有 (3) 的形式. 如果我们能成功地选取积分因子便能做到这一点, 就是说, 选取这样的复值函数 $\lambda = \lambda(p, q)$, 使其满足两个等式

$$\lambda\left(\sqrt{E}dp + \frac{F + i\sqrt{g}}{\sqrt{E}}dq\right) = du + iv,$$

$$\bar{\lambda}\left(\sqrt{E}dp + \frac{F - i\sqrt{g}}{\sqrt{E}}dq\right) = du - iv$$

(注意, 这两个等式的第二个可由第一个取共轭得到). 事实上, 如果找到了这个函数 $\lambda(p, q)$, 则将这两个等式相乘便得到

$$|\lambda|^2 dl^2 = du^2 + dv^2, \quad 即 \quad dl^2 = |\lambda|^{-2}(du^2 + dv^2),$$

从而取 $f(u, v) = |\lambda|^{-2}$. 于是设 $u(p, q), v(p, q), \lambda(p, q)$ 为待定函数. 这些函数应满足方程

$$\lambda\left(\sqrt{E}dp + \frac{F + i\sqrt{g}}{\sqrt{E}}dq\right) = du + idv = \left(\frac{\partial u}{\partial p} + i\frac{\partial v}{\partial p}\right)dp + \left(\frac{\partial u}{\partial q} + i\frac{\partial v}{\partial q}\right)dq.$$

于是

$$\lambda\sqrt{E} = \frac{\partial u}{\partial p} + i\frac{\partial v}{\partial p}, \quad \lambda\frac{F + i\sqrt{g}}{\sqrt{E}} = \frac{\partial u}{\partial q} + i\frac{\partial v}{\partial q}. \tag{*}$$

消去 λ 给出

$$(F + i\sqrt{g})\left(\frac{\partial u}{\partial p} + i\frac{\partial v}{\partial p}\right) = E\left(\frac{\partial u}{\partial q} + i\frac{\partial v}{\partial q}\right)$$

或者

$$F\frac{\partial u}{\partial p} - \sqrt{g}\frac{\partial v}{\partial p} = E\frac{\partial u}{\partial q}, \quad \sqrt{g}\frac{\partial u}{\partial p} + F\frac{\partial v}{\partial p} = E\frac{\partial v}{\partial q}.$$

由此得到

$$
\begin{aligned}
&\text{a)}\ \frac{\partial v}{\partial p}=\frac{F\dfrac{\partial u}{\partial p}-E\dfrac{\partial u}{\partial q}}{\sqrt{g}},\quad \frac{\partial v}{\partial q}=\frac{G\dfrac{\partial u}{\partial p}-F\dfrac{\partial u}{\partial q}}{\sqrt{g}};\\
&\text{b)}\ \frac{\partial u}{\partial p}=-\frac{E\dfrac{\partial v}{\partial q}-F\dfrac{\partial v}{\partial p}}{\sqrt{g}},\quad \frac{\partial u}{\partial q}=\frac{F\dfrac{\partial v}{\partial q}-G\dfrac{\partial v}{\partial p}}{\sqrt{g}}.
\end{aligned}
\tag{**}
$$

由于 $\partial^2_{pq}=\partial^2_{qp}$, 于是由 (**, a) 和 (**, b) 我们得到下面的方程:$Lu=0, Lv=0$, 这里的微分算子 L 是

$$
L=\frac{\partial}{\partial q}\left[\frac{F\dfrac{\partial}{\partial p}-E\dfrac{\partial}{\partial q}}{\sqrt{EG-F^2}}\right]+\frac{\partial}{\partial p}\left[\frac{F\dfrac{\partial}{\partial q}-G\dfrac{\partial}{\partial p}}{\sqrt{EG-F^2}}\right].
$$

称方程 $Lf=0$ 为贝尔特拉米方程, 而算子 L 为贝尔特拉米算子. 因此我们知道了所要找的函数 u, v 应该满足贝尔特拉米方程. 由微分方程理论知道 (自然我们将不给出证明), 如果函数 E, F, G 解析, 则方程 $Lf=0$ 总有解. 于是, 等式 (*), (**) 便确定了函数 u, v, λ. 如果我们令 $w=u+iv, \bar{w}=u-iv$, 则曲面的度量便可写为

$$
dl^2=g(w,\bar{w})dwd\bar{w}.
$$

因此, 我们的曲面可以认为是局部复解析的. □

如在引理 12. 4 所清楚表明的那样, 等温坐标不是唯一决定的. 考虑两组等温坐标 (p,q) 和 (u,v), 并设 $u=u(p,q), v=v(p,q)$.

我们知道, 如果变换保持度量的共形形式不变, 则函数 $w(z,\bar{z})$ 是变量 z 或者 $\bar{z}$ 的复解析函数. 因此, 在等温坐标下保持度量的共形形式不变的变换的全部集合, 它由在各种不同的复解析变换 $w(z)$ 中加进了复共轭 $z\mapsto\bar{z}$ 组成的.

例 13.1. 考虑具复坐标 $z=x+iy$ 的复平面 $\mathbb{C}$. 欧氏度量 $dl^2=dx^2+dy^2$ 具有共形形式, 即

$$
dl^2=dzd\bar{z}. \tag{4}
$$

考虑任意一个分式线性变换

$$
w=\frac{az+b}{cz+d},\quad ad-bc=1. \tag{5}
$$

于是有

$$
dl^2=dzd\bar{z}=|cz+d|^4dwd\bar{w}, \tag{6}
$$

即分式线性变换保持平面的欧氏度量的共形形式.

设在三维欧氏空间中给出了有共形坐标 u, v 的曲面, 使得曲面上的度量为

$$
dl^2=g(u,v)(du^2+dv^2). \tag{7}
$$

我们要推导在共形坐标下的高斯曲率的公式.

定理 13.2. 具度量 (7) 的 $\mathbb{R}^3$ 中曲面的高斯曲率为

$$K = -\frac{1}{2g}\Delta \ln g, \tag{8}$$

其中 $\Delta = \frac{\partial^2}{\partial u^2} + \frac{\partial^2}{\partial v^2}$ 为拉普拉斯算子.

证明 设曲面以参数形式给出:$r = r(u,v), r = (x,y,z)$. 那么方程 (7) 表示

$$\langle r_u, r_u\rangle = \langle r_v, r_v\rangle = g, \quad \langle r_u, r_v\rangle = 0. \tag{9}$$

对这些等式作对 u, v 的微分得到

$$\begin{aligned}
&\frac{1}{2}\frac{\partial g}{\partial u} = \langle r_{uu}, r_u\rangle = \langle r_{uv}, r_v\rangle,\\
&\frac{1}{2}\frac{\partial g}{\partial v} = \langle r_{vv}, r_v\rangle = \langle r_{uv}, r_u\rangle,\\
&\langle r_{uu}, r_v\rangle + \langle r_u, r_{uv}\rangle = 0 = \langle r_{uv}, r_v\rangle + \langle r_u, r_{vv}\rangle.
\end{aligned} \tag{10}$$

令

$$e_1 = \frac{1}{\sqrt{g}} r_u, \quad e_2 = \frac{1}{\sqrt{g}} r_v, \quad n = [e_1, e_2], \tag{11}$$

给出了单位向量 e_1, e_2, n.

在曲面每点上的标架 (e_1, e_2, n) 是法正交的, 这时向量 n 垂直于曲面, 而向量 e_1, e_2 切于它. 按定义, 第二基本型的系数为

$$b_{11} = L = \langle r_{uu}, n\rangle, \quad b_{12} = M = \langle r_{uv}, n\rangle, \quad b_{22} = N = \langle r_{vv}, n\rangle. \tag{12}$$

从公式 (10) 和 (12) 得到向量 r_{uu}, r_{uv}, r_{vv} 在基底 e_1, e_2, n 下的坐标为

$$\begin{aligned}
r_{uu} &= \left(\frac{1}{2\sqrt{g}}\frac{\partial g}{\partial u}, -\frac{1}{2\sqrt{g}}\frac{\partial g}{\partial v}, L\right),\\
r_{uv} &= \left(\frac{1}{2\sqrt{g}}\frac{\partial g}{\partial v}, \frac{1}{2\sqrt{g}}\frac{\partial g}{\partial u}, M\right),\\
r_{vv} &= \left(-\frac{1}{2\sqrt{g}}\frac{\partial g}{\partial u}, \frac{1}{2\sqrt{g}}\frac{\partial g}{\partial v}, N\right).
\end{aligned} \tag{13}$$

因此成立公式

$$\langle r_{uu}, r_{vv}\rangle - \langle r_{uv}, r_{uv}\rangle = LN - M^2 - \frac{1}{2g}\left[\left(\frac{\partial g}{\partial u}\right)^2 + \left(\frac{\partial g}{\partial v}\right)^2\right]. \tag{14}$$

又由公式 (10),(14) 我们得到：

$$\begin{aligned}\frac{1}{2}\frac{\partial^2 g}{\partial u^2} &= \langle r_{uuv}, r_v\rangle + \langle r_{uv}, r_{uv}\rangle \\ &= \frac{\partial}{\partial v}\langle r_{uu}, r_v\rangle - \langle r_{uu}, r_{vv}\rangle + \langle r_{uv}, r_{uv}\rangle \\ &= -\frac{1}{2}\frac{\partial^2 g}{\partial v^2} - (LN - M)^2 + \frac{1}{2g}\left[\left(\frac{\partial g}{\partial u}\right)^2 + \left(\frac{\partial g}{\partial v}\right)^2\right].\end{aligned}$$

那么由 (8.26) 高斯曲率为

$$K = \frac{\det(b_{ij})}{\det(g_{ij})} = \frac{LN - M^2}{g^2} = -\frac{1}{2g}\Delta \ln g.$$

定理得证. □

如果将度量 (7) 写为复形式：

$$dl^2 = h(z,\bar{z})dzd\bar{z},$$

则高斯曲率的公式 (8) 具形式

$$K = -\frac{2}{h}\frac{\partial^2}{\partial z \partial \bar{z}}(\ln h). \tag{15}$$

2. 在共形形式下的球面度量和罗巴切夫斯基平面的度量

在实的情形下，球面的共形坐标已在 §9 中给出. 我们考察半径 $R = 1$ 的球面和它到其赤道平面的球极投影，平面上的坐标为 x, y. 回想在坐标 x, y 下球面度量 dl^2 有形式 (公式 (9.4))

$$dl^2 = 4\frac{dx^2 + dy^2}{(1 + x^2 + y^2)^2} \tag{16}$$

如果 $z = x + iy, \bar{z} = x - iy$，则公式 (16) 有形式

$$dl^2 = \frac{4}{(1 + |z|^2)^2}dzd\bar{z}, \tag{17}$$

其中 z 为复平面 $\mathbb{C}$ 上的坐标.

对于罗巴切夫斯基平面，我们考虑在球极投影下的平面坐标 x, y 下的 dl^2 的形式 ($R = 1$ 时的公式 (10.7))

$$dl^2 = \frac{4}{(1 - (x^2 + y^2))^2}(dx^2 + dy^2). \tag{18}$$

公式 (18) 的复形式可表示为

$$dl^2 = \frac{4}{(1 - |z|^2)^2}dzdz, \quad |z| < 1, \tag{19}$$

其中 z 为单位圆盘 (庞加莱模型的度量)的复坐标. 如果令 $z=\dfrac{1+iw}{1-iw}$ 把单位圆盘映到上半平面 $\mathrm{Im}w>0$, 则罗巴切夫斯基平面的度量为

$$dl^2=\frac{-4}{(w-\bar{w})^2}dwd\bar{w},\quad \mathrm{Im}\,w>0 \tag{20}$$

(克莱因模型的度量).

我们要弄清在复表示下这些度量的运动群是如何构造的. 在上一小节中我们看到分式线性变换的确保持共形形式不变. 我们要在分式变换中找出在我们度量下的运动 (我们将在第四章中证明没有其他的运动). 分式运动设为

$$z=\frac{aw+b}{cw+d},\quad ad-bc=1. \tag{21}$$

对于球面度量我们有

$$dl^2=\frac{4dzd\bar{z}}{(1+|z|^2)^2}=\frac{4dwd\bar{w}}{[|aw+b|^2+|cw+d|^2]^2} \tag{22}$$

$$=\frac{4dwd\bar{w}}{[|b|^2+|d|^2+w(a\bar{b}+c\bar{d})+\bar{w}(\bar{a}b+\bar{c}d)+(|a|^2+|c|^2)|w|^2]^2}.$$

为使变换 (21) 为球面度量 (17) 的运动, 因而它应满足下面等式

$$|b|^2+|d|^2=1,\quad a\bar{b}+c\bar{d}=0,\quad |a|^2+|c|^2=1, \tag{23}$$

同时我们还有

$$ad-bc=1.$$

这样, 矩阵 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 属于 $SU(2)$. 记得在 §11.3 中, 我们在将群 $SL(2,\mathbb{C})$ 等同于所有分式变换的群时应该是模去子群 (±1) 的商群. 因此我们证明了下面的论述.

定理 13.3. 球面 S^2 度量的 (正常) 运动的分式线性群同构于商群 $SU(2)/\pm1$.

推论 存在群同构 $SU(2)/\pm1\simeq SO(3)$.(用球极投影验证!)

注 为了得出整个球面运动群, 我们应该在旋转之中再加入反射. 这对应于在分式线性变换 $z=\dfrac{aw+b}{cw+d}$, $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\in SU(2)$ 再加上复共轭 $z\mapsto\bar{z}$.

转向讨论罗巴切夫斯基平面. 从庞加莱模型开始. 类似于前面那样, 计算表明在作变换 (21) 后得到

$$dl^2=\frac{4dzd\bar{z}}{(1-|z|^2)^2}=\frac{4dwd\bar{w}}{[|d|^2-|b|^2+(c\bar{d}-a\bar{b})w+(\bar{a}b-\bar{c}d)\bar{w}+(|c|^2-|a|^2)|w|^2]^2}.$$

因此应满足条件

$$|d|^2-|b|^2=1,\quad |c|^2-|a|^2=-1,\quad c\bar{d}-a\bar{b}=0, \tag{24}$$

还有 $ad-bc=1$. 于是我们得到矩阵 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 属于群 $SU(1,1)$ 的结论. 注意, 满足了条件 (24) 便保证了在分式线性变换 $z=\dfrac{aw+b}{cw+d}$ 下把单位圆盘变成了自己, 其理由是: 它的边界圆周 $|z|=1$ 变成了圆周 $|w|=1$.

现在考虑克莱因模型. 我们首先寻找能将上半平面 $\mathrm{Im}\ z>0$ 变成自己的形如 (21) 的变换. 为此应满足条件

$$\text{如果 } \mathrm{Im}\ w=0, \quad \text{则} \quad \mathrm{Im}\ z=0. \tag{25}$$

不难看出, 如果 w 为实的而 $z=\dfrac{aw+b}{cw+d}$, 则虚部 $\mathrm{Im}\ z$ 有形式

$$\mathrm{Im}\ z=\frac{w^2\mathrm{Im}(a\bar{c})+w\mathrm{Im}(b\bar{c}+a\bar{d})+\mathrm{Im}(b\bar{d})}{|cw+d|^2}.$$

由于 w 的任意性, 应满足等式

$$\mathrm{Im}(a\bar{c})=\mathrm{Im}(b\bar{c}+a\bar{d})=\mathrm{Im}(b\bar{d})=0.$$

于是从 (25) 所要求的已经得出 a,b,c,d 是实数 (请验证!) 的结论. 具有实 a,b,c,d 的所有变换 (21) 便是在克莱因模型的罗巴切夫斯基度量 $dl^2=\dfrac{4dzd\bar{z}}{-(z-\bar{z})^2}$ 下的运动. 我们不再进行类似于上面各种情形的计算. 令人注目的事实是: 保持由罗巴切夫斯基定义的区域不变的所有分式线性变换, 必自动地表示了运动. 因此证明了下面的论断.

定理 13.4. 罗巴切夫斯基度量的分式线性 (正常的) 运动群同构于

a) 在庞加莱模型时的群 $SU(1,1)/\pm 1$,

b) 在克莱因模型时的群 $SL(2,\mathbb{R})/\pm 1$,

c) 群 $SO(1,2)$(准确地说, 它的连通分支).

定理的最后一个论断出自如下的事实: 群 $SO(1,2)$ 的单位元的连通分支是由保持上半伪球不变的 $\mathbb{R}^3$ 中运动组成 (参看 §10.1).

注 为了得到庞加莱模型的全部运动群还应在分式线性变换上加上复共轭 $z\mapsto \bar{z}$(显然, 它将单位圆盘映为自己, 从而给出度量 (19) 下的运动).

在克莱因模型中应加上变换 $z\mapsto -\bar{z}$, 它把上半平面变到自己, 从而给出度量 (20) 下的运动.

推论 群 $SU(1,1)/\pm 1, SL(2,\mathbb{R})/\pm 1$ 和 $SO(1,2)$ 中的连通分支都相互同构.(以适当坐标变换验证它!)

3. 常曲率曲面

设曲面的度量写为复形式为 $dl^2 = h(z,\bar{z})dzd\bar{z}$. 令 $h = e^{\varphi}$(因为 $h(z,\bar{z}) > 0$, 这是有意义的), 那么高斯曲率便有 (公式 (8),(15)) 形式

$$K = -\frac{1}{2}e^{-\varphi}\Delta\varphi, \quad K = -2e^{-\varphi}\frac{\partial^2\varphi}{\partial z\partial\bar{z}}. \tag{26}$$

如果曲率 K 为常数, 则我们得到了关于函数 φ 的方程 (刘维尔方程):

$$\Delta\varphi = -2Ke^{\varphi}, \quad \frac{\partial^2\varphi}{\partial z\partial\bar{z}} = -\frac{K}{2}e^{\varphi}. \tag{27}$$

定理 13.5. 度量为 $dl^2 = h(z,\bar{z})dzd\bar{z}$ 且有常曲率 K = 常数的曲面 (局部) 等距同构于

a) 在 $K > 0$ 时的球面,

b) 在 $K = 0$ 时的欧氏平面,

c) 在 $K < 0$ 时的罗巴切夫斯基平面.

证明 由 (26) 我们有

$$\begin{aligned} 0 &= \frac{\partial}{\partial z}\left(-\frac{K}{2}\right) = \frac{\partial}{\partial z}\left(e^{-\varphi}\frac{\partial^2\varphi}{\partial z\partial\bar{z}}\right) \\ &= e^{-\varphi}\left(\frac{\partial^3\varphi}{\partial z^2\partial\bar{z}} - \frac{\partial\varphi}{\partial z}\frac{\partial^2\varphi}{\partial z\partial\bar{z}}\right) = e^{-\varphi}\frac{\partial}{\partial\bar{z}}\left(\frac{\partial^2\varphi}{\partial z^2} - \frac{1}{2}\left(\frac{\partial\varphi}{\partial z}\right)^2\right), \end{aligned}$$

由此知 $\dfrac{\partial^2\varphi}{\partial z^2} - \dfrac{1}{2}\left(\dfrac{\partial\varphi}{\partial z}\right)^2 = \psi(z)$, 其中 $\psi(z)$ 是一个解析函数. 作复解析变换 $z = f(w)$, 我们得到

$$h(z,\bar{z}) \to \tilde{h}(w,\bar{w}) = h\left|\frac{df}{dw}\right|^2,$$

$$\varphi \to \tilde{\varphi}(w,\bar{w}) = \varphi(z,\bar{z}) + \ln\frac{df}{dz} + \ln\frac{d\bar{f}}{d\bar{z}}.$$

在新变量下我们也有 $\dfrac{\partial^2\tilde{\varphi}}{\partial w^2} - \dfrac{1}{2}\left(\dfrac{\partial\tilde{\varphi}}{\partial w}\right)^2 = \tilde{\psi}(w)$, 其中 $\tilde{\psi}(w)$ 为 w 的解析函数, 并得它的表达式

$$\tilde{\psi}(w) = \psi(z)(f')^2 + \frac{f'''}{f'} - \frac{3}{2}\left(\frac{f''}{f'}\right)^2.$$

其中 $f' = \dfrac{df}{dw}$. 可以选取函数 f, 使得函数 $\tilde{\psi}(w)$ 变为 0. 为此需要解方程

$$\frac{f'''}{f'} - \frac{3}{2}\left(\frac{f''}{f'}\right)^2 = -\psi(f(w))(f')^2 \tag{28}$$

(在复变函数论中, 此方程的左端被称为 "施瓦茨导数")①.

①本书假定了此方程的可解性.

在进行这个变换后我们得到

$$\frac{\partial^2 e^{-\tilde\varphi/2}}{\partial w^2} = -\frac{1}{2}e^{-\tilde\varphi/2}\left(\frac{\partial^2\tilde\varphi}{\partial w^2} - \frac{1}{2}\left(\frac{\partial\tilde\varphi}{\partial w}\right)^2\right) = 0. \tag{29}$$

由于 $e^{-\tilde\varphi/2}$ 为实的, 我们也有

$$\frac{\partial^2 e^{-\tilde\varphi/2}}{\partial \bar w^2} = 0. \tag{30}$$

由 (29) 和 (30) 得到

$$e^{-\tilde\varphi/2} = aw\bar w + bw + \bar b\bar w + c,$$

其中 a, b 为实的常数, b 为复数. 我们得到度量形式

$$dl^2 = h(z,\bar z)dzd\bar z = \frac{dwd\bar w}{(aw\bar w + bw + \bar b\bar w + c)^2}. \tag{31}$$

这个度量的曲率等于 $K = 4(ac - b\bar b)$(参看公式 (26) 的第一个). 分式线性变换 $w = \dfrac{\alpha z + \beta}{\gamma z + \delta}$ 可以将形式 (31) 转变为

a) $\dfrac{4R^2dzd\bar z}{(1+|z|^2)^2}$, $\quad K = 4(ac - b\bar b) = R^{-2} > 0$;

b) $dzd\bar z$, $\quad K = 4(ac - b\bar b) = 0$

c) $\dfrac{4R^2dzd\bar z}{(1-|z|^2)^2}$, $\quad K = 4(ac - b\bar b) = -R^{-2} < 0$.

我们得到了球面的, 欧氏平面的和罗巴切夫斯基平面的度量. 定理得证. □

习题

13.1. 设度量为

$$dl^2 = dx^2 + f(x)dy^2, \quad 0 < f(x) < \infty.$$

证明此度量可转换为共形形式 $dl^2 = g(u,v)(du^2 + dv^2)$.

§14. 作为 N 维空间中的曲面变换群

1. 在单位元的邻域内的坐标

考虑矩阵群 $GL(n,\mathbb{R})$, 其行列式不为零

$$A = (a^i_j), \quad \det(a^i_j) \neq 0. \tag{1}$$

以 $M(n,\mathbb{R})$ 表示全部矩阵的空间, 则条件 (1) 给出了此空间中的一个区域. 此空间是 n^2 维线性空间. 因此整个线性群是线性空间 $\mathbb{R}^{n^2}$ 中的一个区域. 矩阵的每个分量 a^i_j

作为空间 $M(n,\mathbb{R})$ 的坐标. 如果 $A=(a^i_j), B=(b^i_j)$ 为两个 n 阶矩阵, 则它们的乘积 $C=AB$ 有形式 $C=(c^i_j)$,

$$c^i_j = a^i_k b^k_j, \quad i,j=1,2,\cdots,n. \tag{2}$$

由公式 (2) 知道, 两个矩阵积的坐标通过每个因子的坐标以光滑函数 (它是坐标的多项式) 表达. 换句话说, 乘积规则由直积的光滑映射决定

$$GL(n,\mathbb{R})\times GL(n,\mathbb{R})\to GL(n,\mathbb{R}),$$

其中 $(A,B)\mapsto AB$.

我们在全体 n 阶矩阵的空间 $\mathbb{R}^{n^2}$ 中引进欧氏度量:

$$|A|^2=\sum_{i,j}|a^i_j|^2, \quad A=(a^i_j). \tag{3}$$

显然, 我们有

$$|A+B|\leqslant|A|+|B|. \tag{4}$$

度量 (3) 对于矩阵的积有下面的性质:

引理 14.1. 成立不等式

$$|AB|\leqslant|A||B|. \tag{5}$$

证明 可由下面不等式推导出来

$$\left(\sum x_iy_i\right)^2\leqslant\left(\sum x_i^2\right)\left(\sum y_i^2\right) \tag{6}$$

[不等式 (6) 的推导 : $(\sum x_i^2)(\sum y_i^2)-(\sum x_iy_i)^2=\frac{1}{2}\sum(x_iy_j-x_jy_i)^2$]. 引理证完.□

我们将在单位矩阵 $1\in GL(n,\mathbb{R})$ 的邻域中建立其他方便的坐标系. 考虑在整个矩阵空间中的单位球 (无边缘的)$|X|<1, X=(x^i_j)$.

引理 14.2. 如果 $|X|<1$, 则矩阵 $A=1+X$ 可逆, 即

$$A=1+X\in GL(n,\mathbb{R}).$$

证明 考虑由矩阵构成的级数

$$B=1-X+X^2-X^3+\cdots. \tag{7}$$

我们来证明此级数收敛. 利用不等式 (4) 和 (5), 我们有

$$\begin{aligned}|X^m-X^{m+1}+\cdots\pm X^{m+k-1}|&\leqslant|X^m||1+|X|+\cdots+|X|^{k-1}|\\&=|X|^m\frac{1-|X|^k}{1-|X|}.\end{aligned}$$

于是级数 (7) 的部分和序列在 $|X|<1$ 时是个柯西序列, 从而此级数收敛. 同时

$$AB=(1+X)(1-X+X^2-X^3+\cdots)=1,\quad B=A^{-1}.$$

引理证完. □

我们在群 $GL(n,\mathbb{R})$ 的单位元的邻域

$$|A-1|<1 \tag{8}$$

中引进坐标. 如果 $A=(a^i_j)$, 则矩阵 A 的坐标 x^i_j 等于

$$x^i_j(A)=a^i_j-\delta^i_j,\quad x^i_j(1)=0, \tag{9}$$

$$\delta^i_j=\begin{cases}1, & i=j,\\ 0, & i\neq j.\end{cases}$$

注 类似地可以在以任意点 $B_0\in GL(n,\mathbb{R})$ 为中心的邻域中引进坐标. 事实上, 如果我们对此邻域中所有矩阵都乘以 B_0^{-1}, 则此邻域便转换为单位元的邻域, 而在那里已经构建了坐标. 形式地, 这个过程可以这样给出: 如果 $C=B_0^{-1}=(c^i_j)$, 则我们令

$$\begin{aligned}y^i_j(A)&=c^i_k a^k_j-\delta^i_j,\\ y^i_j(B_0)&=0.\end{aligned} \tag{10}$$

坐标 y^i_j 对于满足

$$|A-B_0|<|B_0| \tag{11}$$

的矩阵 A 适用.

因此我们对于群 $GL(n,\mathbb{R})$, 在其任意点的邻域内都建立了局部坐标.

回到 $M(n,\mathbb{R})$ 中原始的坐标系 (即用矩阵的分量元), 不难证明在单位元, 群 $GL(n,\mathbb{R})$ 的切空间恒同于所有 n 阶矩阵的空间. 考虑曲线 $A(t)\in GL(n,\mathbb{R})$, 即依赖于参数 t 的矩阵族. 假设此曲线在 $t=0$ 时经过单位元, 即 $A(0)=1$. 于是此曲线在 $t=0$ 时的切向量 (速度向量) 为矩阵 $\dot{A}(t)|_{t=0}$. 反过来说, 设 X 为任意矩阵. 于是曲线 $A(t)=1+tX$ 在 t 充分靠零时必落在 $GL(n,\mathbb{R})$ 中. 显然有

$$A(0)=1,\quad \dot{A}(0)=X;$$

这样便证明了所有在单位元切于群 $GL(n,\mathbb{R})$ 的全体向量的集合完全等同于所有 n 阶矩阵的集合.

在 §4 和 §6 所考虑过的变换群都是在全体矩阵的空间中由方程给出来的. 这样, 行列式等于 1 的 n 阶矩阵群 $SL(n,\mathbb{R})$ 便是由一个方程给出的

$$\det A=1. \tag{12}$$

这是在所有矩阵的空间中的超曲面, 它整个位于 $GL(n,\mathbb{R})$ 中.

定理 14.1. $SL(n,\mathbb{R})$ 是所有矩阵的空间中的非异曲面.

证明 我们先证明 $1\in SL(n,\mathbb{R})$ 是这个超曲面的非异点. 为此只需证明 (参看 §7.2) 在这点的 $SL(n,\mathbb{R})$ 的切空间正好有维数 n^2-1 即可. 事实上, 设 $\det A(t)=1$, 并且 $A(0)=1, \frac{d}{dt}A(t)|_{t=0}=X$. 由行列式的微分法则, 我们有

$$0=\frac{d}{dt}(\det A(t))|_{t=0}=\mathrm{Tr}\ X,$$

其中 Tr 表示迹. 因而 $\mathrm{Tr}\ X=0$ 是 $SL(n,\mathbb{R})$ 在点 1 的切空间的方程 (不要忘记, 雅可比矩阵 $\frac{\partial \det A}{\partial a_j^i}$ 在 $A=1$ 处的分量, 是这个方程的系数, 其中 $A=(a_j^i)$. 因此, 我们证明了群 $SL(n,\mathbb{R})$ 在单位元处的切空间与迹为 0 的矩阵的集合等同. 这样的矩阵的空间其维数等于 n^2-1. 于是此群的点 1 是 $SL(n,\mathbb{R})$ 的非异点. 进一步, 让 B 为群 $SL(n,\mathbb{R})$ 中的任意点 (我们称 $SL(n,\mathbb{R})$ 为么模矩阵群). 对矩阵 B 的邻域中的所有矩阵乘以 B^{-1}. 于是 B 变成了 1, 而它的邻域变成了 1 的邻域. 这个映射是光滑和非退化的, 从而点 B 为非异. 证完. □

基于研究群在单位元的切空间是在这里所使用的方法; 我们也把它应用于其他的矩阵群. 因此我们常常只证明对应曲面在点 1 的非异性.(由定理 14.1 的证明清楚表明这已足够了.)

现在来考察 n 阶正交矩阵群 $O(n)$. 它对应的 $\mathbb{R}^{n^2}$ 中的曲面由一个方程组给出:

$$\sum_k a_k^i a_k^j=\delta^{ij},\quad AA^{\mathrm{T}}=1,\quad A=(a_j^i).\tag{13}$$

在 (13) 的这些方程中显然有所重复, 它们可由指标 i,j 的置换得到. 这样还剩下 $\frac{n(n+1)}{2}$ 个方程. 我们应该指出这个方程组的秩等于 $n^2-\frac{n(n+1)}{2}=\frac{n(n-1)}{2}$. 这表明在群 $O(n)$ 的单位元的切空间的维数等于 $\frac{n(n-1)}{2}$. 我们来证明 $O(n)$ 在单位元的切空间与所有反称矩阵的空间等同. 如果 $A(t)\in O(n), A(0)=1$ 为一族正交矩阵, $X=\frac{d}{dt}A(t)|_{t=0}$, 则 $0=\frac{d}{dt}(A(t)A^{\mathrm{T}}(t))|_{t=0}=X^{\mathrm{T}}+X=0$(参看 §5.3). 这是 $O(n)$ 在单位元处的切空间的方程, 表明这些切向量与反称矩阵的空间相同. 所有反称矩阵的空间的维数显然等于 $\frac{n(n-1)}{2}$(作为此空间的笛卡儿坐标可取为矩阵的分量 x_j^i, 满足 $i<j$). 由此得到曲面 $O(n)$ 的非异性.

特别, 作为 $O(n)$ 的连通分支的群 $SO(n)$ 也是矩阵空间 $\mathbb{R}^{n^2}$ 中的非异曲面.

例 14.1. 考虑三维空间中的旋转群 $SO(3)$. 可以取从解析几何中知道的欧拉角作为这个非异曲面上的局部坐标. 如果旋转将坐标系 (x,y,z) 变为 (x',y',z')(图 17), 则这个旋转可以以一串下面三种形式的旋转组成:

a) 绕 z 轴旋转角 φ. 这时 x 轴变为结点线.①

b) 绕结点线旋转角 θ.z 轴变为 z' 轴.

c) 绕 z' 轴旋转角 ψ. 结点线变为 x' 轴.

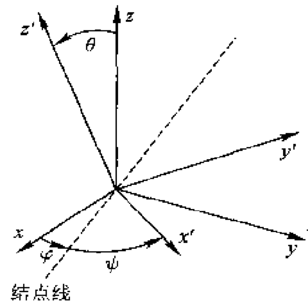


图 17

也可在 $SO(3)$ 上引进其他的局部坐标. 每个旋转可以由决定旋转轴和旋转角 $\varphi, 0\leqslant\varphi\leqslant\pi$ 而给出. 轴可以如此定向: 我们认定绕轴旋转角 φ 是按逆时针方向进行的, 因此每个旋转可以由向量 $\bar{\varphi}$ 给出: 向量 $\bar{\varphi}$ 的方向给出旋转轴, 而模 $|\bar{\varphi}|$ 为旋转角. 那么, 群 $SO(3)$ 的任意点对应于三维空间中半径为 π 的球中的点, 即向量 $\bar{\varphi}$ 的末端点. 由于旋转角 π 和角 $-\pi$ 重合, 则在此球的表面即半径为 π 的球面上, 对径点映成了群 $SO(3)$ 中的同一个旋转.

在这样的坐标系下此群的单位元是坐标原点, 而在单位元的切空间同于三维向量空间.

现在考虑复的情形. 群 $GL(n,\mathbb{C})$ 是全体矩阵的空间 $M(n,\mathbb{C})$ 中的区域, 而 $M(n,\mathbb{C})$ 为 n^2 维复空间 $\mathbb{C}^{n^2}$. 从实的观点看, 这是 $\mathbb{R}^{2n^2}$ 空间中的区域. 群 $SL(n,\mathbb{C})$ 为维数 n^2-1 的复平面 (实维等于 $2n^2-2$). 群 $SL(n,\mathbb{C})$ 在点 1 的切空间同样与迹为零的复矩阵的空间相合.

现在再考虑酉群 $U(n)$. 它在所有 n 阶复矩阵的空间由方程

$$f^{ij}(A)=\sum_{k=1}^{n}a_k^i\bar{a}_k^j=\delta^{ij},\quad A=(a_j^i),\quad 或\ A\bar{A}^{\mathrm{T}}=1. \tag{14}$$

给出.

我们看到函数 $f^{ij}(A)$ 并不是复解析的: $\dfrac{\partial f^{ij}}{\partial\bar{a}_k^j}\neq 0$. 因此群 $U(n)$ 不是复曲面. 我们要证明它是非异的. 每个方程 $f^{ij}=\delta^{ij}$ 在实的意义下给出了两个方程

$$\operatorname{Re} f^{ij}=\delta^{ij},\quad \operatorname{Im} f^{ij}=0.$$

我们注意有等式

$$f^{ij}=\bar{f}^{ji},$$

故方程 $f^{ij}=0$ 和 $f^{ji}=0$ 当 $i\neq j$ 时等价. 另外, $\operatorname{Im} f^{ii}=0$. 于是方程 $f^{ii}=1$ 给出了唯一的矩阵 A 满足的实条件. 从而得到 $n+2\cdot\dfrac{n(n-1)}{2}=n^2$ 个不同的实方程. 需要证明群 $U(n)$ 在单位元的切空间有 (实的) 维数 $2n^2-n^2=n^2$. 这个空间与所有 n

①即 (x,y) 平面与 (x',y') 的交线. — 译注.

阶反埃尔米特矩阵的集合重合; 即如下矩阵

$$\overline{X}^{\mathrm{T}} = -X. \tag{15}$$

事实上, 如果 $A(t) \in U(n), A(0) = 1$, 则

$$A(t)\cdot \bar{A}^{\mathrm{T}}(t) = 1, \quad \left.\frac{dA}{dt}\right|_{t=0} = X, \quad 0 = \left.\frac{d}{dt}(A\bar{A}^{\mathrm{T}})\right|_{t=0} = X + \overline{X}^{\mathrm{T}}.$$

所有反埃尔米特矩阵的空间的维数等于 n^2. 其中的笛卡儿坐标为 $\frac{1}{i}x_k^k, k = 1, \cdots, n$, $\mathrm{Re}\ x_j^i, \mathrm{Im}\ x_j^i$, 其中 $i < j$.

行列式为 1 的酉矩阵的群 $SU(n)$ 也由维数为 n^2-1 的非异实曲面表示. 它在单位元的切空间等同于迹为 0 的所有反埃尔米特矩阵的空间.

例 14.2. 考虑群 $SU(2)$. 我们在 §11.3 中看到, $SU(2)$ 中的矩阵有形式

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}, \quad |a|^2 + |b|^2 = 1.$$

如果 $a = x + iy, b = u + iv$, 则方程 $|a|^2 + |b|^2 = 1$ 给出了以 x, y, u, v 为坐标的四维空间中的三维球面.

2. 矩阵的指数映射

设 T 为群 $GL(n, \mathbb{R})$ 的单位元处的切空间. 构造从此切空间到群本身的映射

$$\exp : T \to GL(n, \mathbb{R}), \quad \exp(0) = 1 \tag{16}$$

为

$$\exp\ X = 1 + \frac{X}{1!} + \frac{X^2}{2!} + \cdots. \tag{17}$$

引理 14.3.

1) 级数 (17) 对所有矩阵 X 均收敛.

2) 如果矩阵 X 和 Y 交换:$XY = YX$, 则

$$\exp(X + Y) = (\exp\ X)(\exp\ Y). \tag{18}$$

3) 矩阵 $A = \exp\ X$ 可逆, 且

$$A^{-1} = \exp(-X). \tag{19}$$

4) $\exp(X^{\mathrm{T}}) = (\exp\ X)^{\mathrm{T}}$.

证明 利用不等式 (4) 和 (5), 对于级数 (17) 我们有

$$\left|\frac{X^m}{m!} + \frac{X^{m+1}}{(m+1)!} + \cdots + \frac{X^{m+k-1}}{(m+k-1)!}\right| < \frac{|X|^m}{m!} + \cdots + \frac{|X|^{m+k-1}}{(m+k-1)!}.$$

但对所有值 $|X|$, 级数 $e^{|X|}$ 收敛. 因此级数 (17) 的部分和序列是柯西序列, 故此级数收敛. 现在证明公式 (18). 利用矩阵 X 和 Y 的可置换性得到

$$\begin{aligned}\exp\ X\exp\ Y &= \left(\sum_{k=0}^{\infty}\frac{X^k}{k!}\right)\left(\sum_{l=0}^{\infty}\frac{Y^l}{l!}\right)\\ &= \sum_{m=0}^{\infty}\frac{1}{m!}\left(\sum_{k+l=m}\frac{m!}{k!\ l!}X^kY^l\right)\\ &= \sum_{m=0}^{\infty}\frac{1}{m!}(X+Y)^m = \exp(X+Y).\end{aligned}$$

公式 (19) 由 (18) 得到, 这是因为矩阵 X 和 $-X=Y$ 可交换, 且 $\exp(0)=1$. 论断 4) 显然成立.

引理得证. □

设 G 为上面考虑过的矩阵群 ($G=GL(n,\mathbb{R}), O(n), U(n)$ 等等) 中的某一个. 又设 T 为群 G 在单位元处的切空间. 我们要证明映射 exp 将 T 映到 G 内.

引理 14.4.

1) 如果 $G=SL(n,\mathbb{R}), X\in T$, 即 $\mathrm{Tr}\ X=0$, 则 $A=\exp\ X\in SL(n,\mathbb{R})$, 即 $\det A=1$.

2) 如果 $G=O(n), X\in T$, 即 X 为反称矩阵, 则 $A=\exp\ X\in O(n)$, 即 A 为正交矩阵.

3) 如果 $G=U(n), X\in T$, 即 X 为反埃尔米特矩阵, 则 $A=\exp\ X\in U(n)$, 即矩阵 A 为酉矩阵.

证明 设 $\mathrm{Tr}\ X=0$. 考虑矩阵族

$$A(t)=\exp(tA),$$

t 为参数. 由引理 14. 3 我们有

$$A(t_1+t_2)=A(t_1)A(t_2)$$

(矩阵 t_1X 与 t_2X 可交换). 如果 $f(t)=\det\ A(t)=\det\exp(tX)$, 则 $f(t_1+t_2)=f(t_1)f(t_2)$. 于是, $f(t)=e^{ct}, c$ 为常数. 证 $c=0$: 我们有 $f(t)=\det\exp(tX)=\det(1+tX+o(t))=1+t\mathrm{Tr}\ X+o(t)$. 如果 $\mathrm{Tr}\ X=0$, 则 $c=\frac{df}{dt}\big|_{t=0}=0$. 因此 $\det\ A(t)=1$; 特别 $\det\ A=1$. 论断 1) 得证.

设矩阵 X 为反对称:

$$X^{\mathrm{T}}=-X.$$

于是矩阵 X 和 X^{T} 相互可换. 设 $A=\exp\ X$. 于是由于引理 14.3 我们有

$$AA^{\mathrm{T}}=(\exp\ X)(\exp\ X)^{\mathrm{T}}=\exp(X+X^{\mathrm{T}})=1,\quad A\in O(n).$$

论断 2) 得证.

现在设矩阵 X 为反埃尔米特的

$$X^{\mathrm{T}} = -X$$

(其中上面的横线表示复共轭). 设 $A = \exp\ X$. 于是又由引理 14.3 得到

$$A\bar{A}^{\mathrm{T}} = \exp\ X\overline{\exp\ X}^{\mathrm{T}} = \exp\ X \exp \bar{X}^{\mathrm{T}} = \exp(X + \bar{X}^{\mathrm{T}}) = 1, \quad A \in U(n).$$

引理证完. □

由引理 14.4 也可得到结论：对于群 $G = SO(n), SU(n)$ 来说, 如果矩阵 X 在 T 中, 则指数 $\exp\ X$ 也在 G 中.

引理 14.5. 在坐标原点 (即 $M(n,\mathbb{R})$ 中的零矩阵) 的某个邻域中映射 $\exp\ X$ 是一一对应的.

证明 只需验证映射 exp 在点 O(坐标原点) 的雅可比不为零即可. 设 $A = (a^i_j) = \exp\ X, X = (x^i_j)$. 我们有

$$\frac{\partial a^k_l}{\partial x^i_j} = \delta^k_i \delta^j_l + \cdots,$$

其中的 "$\cdots$" 表示那些在 $X = 0$ 时变成 0 的项. 因此映射 exp 的雅可比矩阵具有形式 $\left(\frac{\partial a^k_l}{\partial x^i_j}\right)\Big|_{x=0} = (\delta^k_i \delta^j_l)$. 这是 n^2 阶的单位矩阵, 从而引理得证. □

映射 exp 让我们可以在群的单位元的邻域中引进方便的局部坐标：以切空间 T 零点的邻域中的笛卡儿坐标 x^i_j 作为这里的坐标. 对坐标 x^i_j 的显式公式为

$$x^i_j(A) = (\ln\ A)^i_j = \left[(A-1) - \frac{(A-1)^2}{2} + \frac{(A-1)^3}{3} - \cdots\right]^i_j. \tag{20}$$

但整个映射 exp 可能不是一一对应的 (甚至也可能没有映满整个群 G; 参看习题 14.3).

注 考虑空间 T 中的通过坐标原点的直线, 即矩阵族 tX, t 为参数. 那么矩阵族 $A(t) = \exp(tX)$ 形成一个单参数子群.

$$\begin{aligned} &A(s)A(t) = A(s+t), \quad A(0) = 1, \\ &A(-t) = A^{-1}(t). \end{aligned} \tag{21}$$

例 14.3. 在群 $SO(3)$ 中显然有绕任意轴旋转的单参数子群. 我们注意到有 $A(t + 2\pi) = A(t)$. 对于绕 z 轴的旋转有

$$A(t) = \begin{pmatrix} \cos\ t & \sin\ t & 0 \\ -\sin\ t & \cos\ t & 0 \\ 0 & 0 & 1 \end{pmatrix} = \exp\left[t\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\right].$$

3. 四元数

四元数的集合是形如下面的实系数的线性组合的全体集合 $\mathbb{H}$

$$q \in \mathbb{H}, \quad q = a + bi + cj + dk, \tag{22}$$

其中 i, j, k 是一些线性独立的符号, 在 $\mathbb{H}$ 中引入双线性乘积为

$$ij = k = -ji, \quad jk = i = -kj, \quad ki = j = -ik, \quad i^2 = j^2 = k^2 = -1, \tag{23}$$

而 $b = c = d = 0$ 的四元数与所有其他四元数可交换. 容易验证具这样定义的乘法下的 $\mathbb{H}$ 成为一个可结合的实数域上的代数, 但不是交换的. 这个代数可以以矩阵形式表示.

引理 14.6. 对于每个四元数 $q = a + bi + cj + dk$ 我们指定一个矩阵 $A(q) \in M(2, \mathbb{C})$, 其中

$$A(q) = \begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix}, \tag{24}$$

(其中 i 为虚单位). 于是有

$$A(q_1 q_2) = A(q_1) A(q_2). \tag{25}$$

注 公式 (25) 表明映射 $q \mapsto A(q)$ 为同态.

证明 只要对 $q = i, j, k$ 导出 (25) 即可. 我们有

$$A(i) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad A(j) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad A(k) = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \tag{26}$$

现在不难推导出

$$A(i)A(j) = A(k),$$

等等. □

注 常常称矩阵

$$\sigma_x = -iA(k), \quad \sigma_y = -iA(j), \quad \sigma_z = -iA(i) \tag{27}$$

为泡利 (Pauli) 矩阵. 这些矩阵满足

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1, \quad \sigma_x \sigma_y = -\sigma_y \sigma_x = i\sigma_z, \cdots. \tag{28}$$

我们在 $\mathbb{H}$ 中引入共轭运算

$$\bar{q} = a - bi - cj - dk, \tag{29}$$

其中 $q = a + bi + cj + dk$.

引理 14.7. 映射 $q \mapsto \bar{q}$ 为代数 $\mathbb{H}$ 的反自同态, 即

$$\overline{q_1 + q_2} = \bar{q}_1 + \bar{q}_2; \quad \overline{q_1 q_2} = \bar{q}_2\, \bar{q}_1. \tag{30}$$

证明 立即可由

$$A(\bar{q}) = \bar{A}^{\mathrm{T}}(q) \tag{31}$$

和映射 $A \mapsto \bar{A}^{\mathrm{T}}$ 为矩阵代数 $M(2,\mathbb{C})$ 的反自同态得到结果. □

我们定义四元数的模 $|q|^2$ 为

$$|q|^2 = q\bar{q} = a^2 + b^2 + c^2 + d^2, \tag{32}$$

其中 $q = a + bi + cj + dk$.

引理 14.8. 四元数全体 $\mathbb{H}$ 构成一个"除环"("非交换域"). 它表示对每个非零四元数 $q, |q|^2 \neq 0$, 有一个确定的逆四元数 q^{-1}, 使得

$$qq^{-1} = 1.$$

证明 可令

$$q^{-1} = \frac{\bar{q}}{|q|^2}, \tag{33}$$

于是 $qq^{-1} = q\bar{q}/|q|^2 = 1$. □

直接计算指出

$$|q|^2 = \det A(q); \tag{34}$$

因而模有性质

$$|q_1 q_2|^2 = |q_1|^2 |q_2|^2. \tag{35}$$

模为 1 的四元数的全体记做 $\mathbb{H}_1$. 由 (35) 知此 $\mathbb{H}_1$ 在乘法下为群, 又由 (33) 得到: 如果 $q \in \mathbb{H}_1$ 则 $q^{-1} = \bar{q}$.

在具坐标 (a, b, c, d) 的四维空间 $\mathbb{R}^4$ 中, $\mathbb{H}_1$ 是个超曲面, 它满足方程

$$a^2 + b^2 + c^2 + d^2 = |q|^2 = 1. \tag{36}$$

因此 $\mathbb{H}_1$ 与 $\mathbb{R}^4$ 中的单位三维球面等同. 我们还注意到, 如果 $q = a + bi + cj + dk, 1 = |q|^2$, 且 $x = a + bi, y = c + di$, 则 $|x|^2 + |y|^2 = 1$, 且矩阵 $A(q)$ 有形式

$$A(q) = \begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix} = \begin{pmatrix} x & y \\ -\bar{y} & \bar{x} \end{pmatrix}. \tag{37}$$

于是 (参看 §11.3) 群 $\mathbb{H}_1$ 同构于群 $SU(2)$.

我们在 §13.2 中看到球面旋转群 $SO(3)$ 同构于商群:$SO(3) \simeq SU(2)/\pm 1$. 我们将再给出一个方法来证明它.

设 $\mathbb{H}_0$ 为满足 $\bar{x} = -x$ 的四元数 x 的三维空间. 在此空间中用公式 $|x|^2 = x\bar{x} = -x^2$ 给出度量. 显然这个空间是欧几里得的.

引理 14.9. 如果 $|q|^2=1$, 则变换

$$\alpha_q : x \mapsto qxq^{-1}, \quad x\in \mathbb{H}_0, \tag{38}$$

为三维欧氏空间 $\mathbb{H}_0=\mathbb{R}^3$ 的旋转.

证明 因为 $\bar{x}=-x, \bar{q}=q^{-1}$, 故而 $\overline{qxq^{-1}}=\bar{q}^{-1}\bar{x}q=-qxq^{-1}$. 这样, 在变换 α_q 下三维空间 $\mathbb{H}_0$ 变到自己. 另外, $|qxq^{-1}|=|x|$, 即其保持 $\mathbb{H}_0$ 中向量长度不变. 引理证完. □

于是, 映射 $q\mapsto \alpha_q$ 为群 $\mathbb{H}_1\simeq SU(2)$ 到三维空间 $\mathbb{R}^3$ 的旋转群的同态. 容易证明, 这个同态的像为整个旋转群, 以及 $\alpha_{-q}=\alpha_q$, 还有事实: 如果 $\alpha_{q_1}=\alpha_{q_2}$, 则 $q_1=\pm q_2$. 由此得到了 $SO(3)\simeq SU(2)/\pm 1$.

再举个例子. 我们要建立群同构 $SU(2)\times SU(2)/\pm 1\simeq SO(4)$. 设 $p,q\in \mathbb{H}_1\simeq SU(2)$. 那么映射

$$\alpha_{p,q} : x\mapsto pxq^{-1}, \quad x\in \mathbb{H}=\mathbb{R}^4, \tag{39}$$

保持四元数 x 的模平方不变, 从而是四维空间 $\mathbb{H}$ 的正交变换 (使坐标原点保持不动). 那么, 映射 $(p,q)\mapsto \alpha_{p,q}$ 是 $SU(2)\times SU(2)$ 到 $O(4)$ 的同态. 因为这个映射使核由偶对 $(1,1),(-1,-1)$(因 $\alpha_{-p,-q}=\alpha_{p,q}$) 组成, 故有单同态 $SU(2)\times SU(2)/\pm 1\to O(4)$. 由于群 $SU(2)\times SU(2)/\pm 1$ 的连通性, 映射 $(p,q)\mapsto \alpha_{p,q}$ 的像在群 $SO(4)$ 中 (在连续映射 (39) 下连通集的像集为连通集).

现在考虑 n 维四元数空间 $\mathbb{H}^n$, 其基为 $e_1,\cdots,e_n$, 四元数坐标为 $q^1,\cdots,q^n$. 有一个重要的事实: 任意四元数 $q=a+bi+cj+dk$ 可以表示为如下形式:

$$q=x+yj=x+j\bar{y}, \quad x=a+bi, \quad y=c+di, \tag{40}$$

其中 x 和 y 可以简单地看作是复数. 如果 $u+vj$ 为另一个四元数, 则有下面等式:

$$(x+yj)(u+vj)=(x+yj)(u+j\bar{v})=(xu-y\bar{v})+(xv+y\bar{u})j. \tag{41}$$

现在空间 $\mathbb{H}^n$ 可以看作为 $2n$ 维复空间 $\mathbb{C}^{2n}$, 其基为 $e_1,\cdots,e_n,je_1,\cdots,je_n$; 复坐标为 $x^1,\cdots,x^n,y^1,\cdots,y^n$, 其中 $q^k=x^k+y^kj$. 事实上, 利用表达式 (40) 我们得到

$$q^ke_k=x^ke_k+y^kje_k=x^ke_k+y^k(je_k). \tag{42}$$

设 $GL(n,\mathbb{H})$ 为空间 $\mathbb{H}^n$ 的可逆线性变换群. 每个变换 $A\in GL(n,\mathbb{H})$ 由矩阵 (λ_l^k) 给出, 其中 $\lambda_l^k\in\mathbb{H}, k,l=1,\cdots,n$, 这时四元数坐标 $q^1,\cdots,q^n$ 按规律

$$q^k\mapsto q^l\lambda_l^k=q'^k \tag{43}$$

变换 (这里因子的次序是很重要的). 利用等式 (41) 我们得到复坐标变换的规则:

$$x^k\mapsto (x^la_l^k-y^l\bar{b}_l^k)=x'^k,$$

$$y^k \mapsto (x^l b_l^k + y^l \bar{a}_l^k) = y'^k, \tag{44}$$

$$\lambda_l^k = a_l^k + b_l^k j.$$

因此, 空间 $\mathbb{H}^n$ 的每个 $\mathbb{H}$ 线性变换由相应的空间 $\mathbb{C}^{2n}$ 的复线性变换给出. 我们得到了群同态

$$c: GL(n, \mathbb{H}) \to GL(2n, \mathbb{C}), \tag{45}$$

同时由 (44) 知道, 如果 $\Lambda = A + Bj$, 则

$$c(\Lambda) = \begin{pmatrix} A & B \\ -\bar{B} & \bar{A} \end{pmatrix}. \tag{46}$$

注 类似于 §11.1, 同态 c 的像可以确定为所有与乘以 j 的算子 J 交换的复矩阵, 即 $J = c(j)$.

我们令

$$|\xi|^2 = \sum_{k=1}^{n} |q^k|^2 = \sum_{k=1}^{n} q^k \bar{q}^k, \quad \xi = q^k e_k \tag{47}$$

来定义 $\mathbb{H}^n$ 中的二次型 $|\xi|^2$. 如果 $q = x + yj$, 则 $|q|^2 = |x|^2 + |y|^2$, 在复坐标 (x^k, y^k) 下 $q^k = x^k + y^k j$, 那么形式 (47) 便等同于标准的向量模的平方

$$\sum |q^k|^2 = \sum |x^k|^2 + \sum |y^k|^2. \tag{48}$$

对应的 $\mathbb{H}^n$ 中的类似的埃尔米特形式便有形式

$$\langle \xi_1, \xi_2 \rangle_{\mathbb{H}} = \sum_{k=1}^{n} q_1^k \bar{q}_2^k, \quad \xi_1 = q_1^k e_k, \quad \xi_2 = q_2^k e_k. \tag{49}$$

定义 14.1. 称所有保持形式 (49) 不变的四元数变换的集合为辛群, 记为 $Sp(n)$. 在复坐标下形式 (49) 表达为

$$\begin{aligned} \langle \xi_1, \xi_2 \rangle_{\mathbb{H}} &= \sum_{k=1}^{n} (x_1^k + y_1^k j)\overline{(x_2^k + y_2^k j)} \\ &= \sum_{k} (x_1^k \bar{x}_2^k + y_1^k \bar{y}_2^k) + \sum_{k} (y_1^k x_2^k - x_1^k y_2^k) j. \end{aligned} \tag{50}$$

设 $A \in Sp(n)$. 由公式 (50) 知道 A 保持埃尔米特形式 $\sum\limits_{k} (x_1^k \bar{x}_2^k + y_1^k \bar{y}_2^k) = \langle \xi_1, \xi_2 \rangle_{\mathbb{C}}$ 不变和反称形式 $\sum\limits_{k} (y_1^k x_2^k - x_1^k y_2^k)$ 不变. 于是 $c(Sp(n))$ 包含在 $U(2n)$ 中由空间 $\mathbb{C}^{2n}$ 中保持反称形式 $\sum\limits_{k} (y_1^k x_2^k - x_1^k y_2^k)$ 不变的酉变换组成.

例 14.4. 群 $Sp(1)$ 同构于 $SU(2)$. 事实上 $c(Sp(1)) \subset U(2)$, 此时 $c(Sp(1))$ 中的矩阵保持了 $y_1x_2 - x_1y_2$ 不变, 即保持了复平行四边形的"面积", 此四边形由向量 (x_1, y_1) 和 (x_2, y_2) 张成. 由于 $c(Sp(1))$ 的矩阵有形式 $\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$, $|a|^2 + |b|^2 = 1$, 我们所得到的这个结果与上面得到的一样. 这是 $SU(2)$ 中矩阵的同一形式 (见 §11.2).

习题

14.1. 证明群 $GL(n, \mathbb{C}), U(n), Sp(n)$ 的连通性 (参看 §4).

14.2. 证明对任意矩阵成立公式: $\det(\exp X) = e^{\mathrm{Tr}\, X}$.

14.3. 证明对 $SL(2, \mathbb{R})$ 指数映射的像不是整个群.

14.4. 描述群 $U(n), SU(n), SO(p,q), SU(p,q)$ 在单位元的切空间的结构.

14.5. 找出 $SL(2, \mathbb{R})$ 中所有单参数子群.

§15. 高维欧氏空间和伪欧氏空间的共形变换

设在 $\mathbb{R}^n$ 中的区域 U 上给出了两个度量 $g_{\alpha\beta}$ 和 $g'_{\alpha\beta}$; 空间的坐标为 $x^1, \cdots, x^n$. 如果 $g'_{\alpha\beta}(x) = \lambda(x) g_{\alpha\beta}(x), \lambda(x) \neq 0$, 我们则说这两个矩阵定义了同一个共形结构. 更一般地, 等价于给定度量 $g_{\alpha\beta}$ 的度量, 用坐标变换 $x = x(y)$, 就得到度量 $g_{\alpha\beta}$ 乘以数值函数 $\lambda(x)$.

定义 15.1. 具度量 $g_{\alpha\beta}$ 的区域 U 到自身的映射 φ: $U \to U$ 被称做*共形度量*是说, 如果度量 $g'_{\alpha\beta} = \dfrac{\partial x^k}{\partial y^\alpha} g_{kl} \dfrac{\partial x^l}{\partial y^\beta}$ 与原来的度量成比例

$$g'_{\alpha\beta} = \lambda g_{\alpha\beta}, \quad \lambda = \lambda(x).$$

进一步考虑 $\mathbb{R}^n = \mathbb{R}^n_{p,q}$ 中 (p,q) 型度量的伪欧氏空间, 其度量为 $g_{\alpha\beta} = g_{\alpha\alpha}\delta_{\alpha\beta}$, $g_{\alpha\alpha} = \pm 1$. 我们特别感兴趣的是具正定度量的欧氏空间 $\mathbb{R}^n$ 和闵可夫斯基空间 $\mathbb{R}^n_1 = \mathbb{R}^n_{1,n-1}$.

这些空间的共形变换里有线性的变换. 显然, 线性的共形变换可归结为下列的类别 (参看 §4,6):

a) 运动: 群 $O(p,q), \lambda(x) \equiv 1$;

b) 伸缩,

$$x \mapsto \lambda x, \quad \lambda = \text{常数};$$

c) 运动和伸缩的组合,

$$x \mapsto \lambda A(x), \quad A \in O(p,q), \quad \lambda = \text{常数}.$$

除此之外还有平移 (也是运动)

$$x \mapsto x + x_0.$$

这是些“显然的”共形变换. 在它们之外还有被称为反演的变换:

d) 反演,

$$x^\alpha \mapsto \frac{x^\alpha - x_0^\alpha}{\langle x - x_0, x - x_0 \rangle}, \tag{1}$$

其中 $\langle , \rangle$ 为内积

$$\langle x - x_0, x - x_0 \rangle = (x^\alpha - x_0^\alpha)(x^\beta - x_0^\beta) g_{\alpha\beta}(x).$$

我们应留意, 映射 (1) 在点 x_0 没有定义; 为了消除这个不确定性, 可以在 $\mathbb{R}^n$ 上添加无穷远处的点, 作为点 x_0 的像.

我们断言反演确是共形变换. 为此我们以 y 表示 x 在反演下的像. 令 $g_{\alpha\beta} = \delta_{\alpha\beta}, x \neq x_0$, 我们得到

$$\sum_{\beta=1}^{n} (y^\beta)^2 = \langle x - x_0, x - x_0 \rangle^{-1}.$$

于是

$$x^\alpha = \frac{y^\alpha}{\sum (y^\beta)^2} + x_0^\alpha.$$

由此计算 $\dfrac{\partial x^k}{\partial y^\alpha}$, 得到了 $g'_{\alpha\beta}$ 的表达式, 它的共形因子为

$$\lambda(x) = \langle x - x_0, x - x_0 \rangle^2. \tag{2}$$

当 $n = 1$ 时所有变换为共形的.

当 $n = 2$ 时有许多共形变换, 它们由任意一个复解析函数所定义 (参看引理 12.4).

当 $n > 2$ 时的情形则不一样. 有定理:

定理 15.1. **(刘维尔)** 任意光滑 (要求所给映射的函数至少有四阶的连续导数[①]) 共形变换, 如果定义在维数 $n \geqslant 3$ 的空间的欧氏 (伪欧氏) 区域上, 则必是运动, 伸缩和反演的复合.

证明 我们给出 $n = 3$ 情形的证明.

在证明中欧氏和伪欧氏的情形在本质上没有差别. 故而我们将只证通常欧氏度量的情形. 设 $y^\alpha = y^\alpha(x^1, \cdots, x^n)$, 为从区域 $U \subset \mathbb{R}^n$ 到区域 $V \subset \mathbb{R}^n$ 的映射. (可以将此映射当作为在区域 U 中引进了新的坐标.) 不失一般性, 可以设点 $O = (0, \cdots, 0)$ 在两个区域 U, V 中且 $y^\alpha(O) = O$. 共形性意味着矩阵 $A = (a_\beta^\alpha) = \left(\dfrac{\partial y^\alpha}{\partial x^\beta}\right)$ 是在

[①] 实际上定理在更少的光滑性下也成立, 但我们不用它.

每点 $x \in U$ 上的线性共形变换的矩阵. 换句话说, 对于 $dx = (dx^\alpha), dy = Adx$ 或 $dy^\alpha = a^\alpha_\beta dx^\beta$, 我们有

$$|dy| = \lambda(x)|dx|. \tag{3}$$

设 η_1, η_2, η_3 为三个向量, 它们两两正交:$\langle \eta_i, \eta_j \rangle = 0, i \neq j$. 我们于是有

$$\langle \eta_i, \eta_j \rangle = 0 = \langle A\eta_i, A\eta_j \rangle, \quad i \neq j, \tag{4}$$

它对所有点 $x \in U$ 成立, 而且 $A = A(x) = \left(\dfrac{\partial y^\alpha}{\partial x^\beta}\right)$. 对 (4) 作对第三个向量 $\eta_k, i \neq k, j \neq k$ 的方向微分, 我们得到

$$\begin{aligned} 0 &= \eta_k^\gamma \frac{\partial}{\partial x^\gamma} \langle A\eta_i, A\eta_j \rangle \\ &= \left\langle \eta_k^\gamma \frac{\partial^2 y}{\partial x^\beta \partial x^\gamma} \eta_i^\beta, A\eta_j \right\rangle + \left\langle A\eta_i, \eta_k^\gamma \frac{\partial^2 y}{\partial x^\beta \partial x^\gamma} \eta_j^\beta \right\rangle. \end{aligned} \tag{5}$$

循环交换指数 (i,j,k), 我们由 (5) 得出三个等式. 从其中两个联立计算第三个, 我们得到了

$$\left\langle \frac{\partial^2 y}{\partial x^\gamma \partial x^\beta} \eta_1^\gamma \eta_2^\beta, A\eta_3 \right\rangle = 0 \tag{6}$$

(可对指标 1,2,3 进行置换).

因此, 如果向量 η_3 正交于向量 η_1 和 η_2, 则向量 $\dfrac{\partial^2 y}{\partial x^\gamma \partial x^\beta} \eta_1^\gamma \eta_2^\beta$ 正交于向量 $A\eta_3$. 由于 $n = 3$, 我们便有了

$$\frac{\partial^2 y}{\partial x^\gamma \partial x^\beta} \eta_1^\gamma \eta_2^\beta = \mu(x)(A\eta_1) + \nu(x)(A\eta_2). \tag{7}$$

由定义知系数 μ 和 ν 有形式

$$\begin{aligned} \mu &= \frac{1}{|A\eta_1|^2} \left\langle \frac{\partial^2 y}{\partial x^\gamma \partial x^\beta} \eta_1^\gamma \eta_2^\beta, A\eta_1 \right\rangle, \\ \nu &= \frac{1}{|A\eta_2|^2} \left\langle \frac{\partial^2 y}{\partial x^\gamma \partial x^\beta} \eta_1^\gamma \eta_2^\beta, A\eta_2 \right\rangle. \end{aligned} \tag{8}$$

另外, 由 (3) 知 $|A\eta_1|^2 = \lambda^2(x)|\eta_1|^2$, 因而

$$\begin{aligned} \mu &= \frac{1}{\lambda^2 |\eta_1|^2} \left[\frac{1}{2} \eta_2^\alpha \frac{\partial}{\partial x^\alpha} \langle A\eta_1, A\eta_1 \rangle \right] = \frac{1}{\lambda} \eta_2^\alpha \frac{\partial \lambda}{\partial x^\alpha}, \\ \nu &= \frac{1}{\lambda} \eta_1^\alpha \frac{\partial \lambda}{\partial x^\alpha}. \end{aligned} \tag{9}$$

以 $\lambda(x)$ 除 (9) 并利用 (8), 我们得到

$$\frac{\partial^2}{\partial x^\gamma \partial x^\beta}(\rho y \eta_1^\gamma \eta_2^\beta) = \left(\frac{\partial^2 \rho}{\partial x^\gamma \partial x^\beta}\eta_1^\gamma \eta_2^\beta\right) y, \tag{10}$$

其中 $\rho = \sqrt{\lambda}$. 按第三个向量 η_3 的方向微分 (10) 得到

$$\frac{\partial^3(\rho y)}{\partial x^\gamma \partial x^\beta \partial x^\delta}\eta_1^\gamma \eta_2^\beta \eta_3^\delta = \left(\frac{\partial^2 \rho}{\partial x^\gamma \partial x^\beta}\eta_1^\gamma \eta_2^\beta\right) A\eta_3 + \left(\frac{\partial^3 \rho}{\partial x^\delta \partial x^\gamma \partial x^\beta}\eta_1^\gamma \eta_2^\beta \eta_3^\delta\right) y. \tag{11}$$

这个等式两端的项对于 (η_1, η_2, η_3) 的置换对称. 因而中间一项应该是对称的:

$$\left(\frac{\partial^2 \rho}{\partial x^\gamma \partial x^\beta}\eta_1^\gamma \eta_2^\beta\right) A\eta_3 = \left(\frac{\partial^2 \rho}{\partial x^\gamma \partial x^\beta}\eta_2^\gamma \eta_3^\beta\right) A\eta_1. \tag{12}$$

因为 $A\eta_3 \neq 0, A\eta_1 \neq 0$ 和 $A\eta_2 \neq 0$(我们设定了从坐标 x 到坐标 y 的变换的雅可比不为零), 得到了

$$\frac{\partial^2 \rho}{\partial x^\gamma \partial x^\beta}\eta_1^\gamma \eta_2^\beta \equiv 0. \tag{13}$$

因此, 在任意一对正交向量上双线性形式 (13) 化为零. 于是它的矩阵以一个比例因子 $\sigma(x)$ 正比于某个度量 $g_{\gamma\delta}$

$$\frac{\partial^2 \rho}{\partial x^\gamma \partial x^\delta} = \sigma(x) g_{\gamma\delta}. \tag{14}$$

我们要证明 $\sigma(x) =$ 常数. 设 ξ_i 为任意一个向量. 微分 (14) 我们有

$$\frac{\partial^3 \rho}{\partial x^\gamma \partial x^\beta \partial x^\delta}\xi_1^\gamma \xi_2^\beta \xi_3^\delta = \left(\frac{\partial \sigma}{\partial x^\gamma}\xi_1^\gamma\right)\langle \xi_2, \xi_3\rangle. \tag{15}$$

互换 ξ_1 和 ξ_2, 并将得到的等式相减便有

$$\left\langle \left(\frac{\partial \sigma}{\partial x^\alpha}\xi_1^\alpha\right)\xi_2 - \left(\frac{\partial \sigma}{\partial x^\gamma}\xi_2^\gamma\right)\xi_1, \xi_3 \right\rangle = 0. \tag{16}$$

由对所有向量 ξ_i 的这些等式便得到 $\sigma =$ 常数. 最后我们有

$$\frac{\partial^2 \rho}{\partial x^\gamma \partial x^\beta} = \sigma g_{\gamma\beta}, \tag{17}$$

$$\sigma = \text{常数}, \quad \rho = \sqrt{\lambda(x)} = a_1|x - x_0|^2 + b_1, \quad a_1, b_1 \text{为常数}.$$

对于逆映射 $x = x(y)$, 同样是共形的 (共形因子为 $\frac{1}{\lambda(x)}$), 我们有

$$\frac{1}{\sqrt{\lambda(x)}} = a_2|y - y_0|^2 + b_2, \quad a_2, b_2 = \text{常数}. \tag{18}$$

由 (17) 和 (18) 得出

$$(a_1|x-x_0|^2+b_1)(a_2|y-y_0|^2+b_2)=1. \tag{19}$$

因为 $y(0)=0$, 故 $x_0=y_0=0$. 由此得到共形映射把任意球面变成球面.

考虑球面族 $|x|=R$. 因为 $y(0)=0$, 射线 $(0,y)$ 变到射线 $(0,x)$.(它与共形映射保持曲线之间夹角不变相关: 垂直于中心在坐标原点的 U 中球面上每点的线段 (射线) 变成了 V 中的曲线, 它也垂直于整个球面.) 另外, 如果以参数 τ 使线段 $(0,y)$ 参数化, $0\leqslant\tau\leqslant t$, 则

$$|y|=\int_0^t\sqrt{\sum_\alpha\left(\frac{dy^\alpha}{d\tau}\right)^2}d\tau=\int_0^t\sqrt{\frac{1}{\lambda}\sum_\alpha\left(\frac{dx^\alpha}{d\tau}\right)^2}d\tau.$$

选取 τ 为线段 $(0,x)$ 上的自然参数, 我们有

$$|y|=\int_0^x\frac{d\tau}{\sqrt{\lambda}}=\int_0^{|x|}\frac{d\tau}{a_1\tau^2+b_1}. \tag{20}$$

如果 $a_1\neq0$ 和 $b_1\neq0$, 则这个积分给出了关于上限的超越函数. 它与 (19) 矛盾. 这表明或是 $a_1=0$, 或是 $b_1=0$.

$a_1=0$ 的情形. 于是 $\lambda=$ 常数. 这是运动和伸缩的结合.

$b_1=0$ 的情形 (这时 $a_1\neq0$). 运用 $x^*=x/|x|^2$ 我们把这个情形化为上面的情形, 因为我们有 $|x^*|=1/|x|$.

对于欧几里得度量及 $n=3$, 定理得证. □

留意一下, 在证明中我们并没有用到度量的欧几里得性质. 因此定理对 $n=3$ 时的伪欧氏度量也成立. 对于 $n>3$ 情形的证明, 其变动并不大. 维数只用于三个正交向量 (η_1,η_2,η_3) 的等式 (7).

习题 15.1. 证明对维数 $n\geqslant3$ 时的 (7).

有了这个习题后, 定理便由进一步计算自动地对于任意的 n 都成立.

我们对欧氏或非欧氏空间中的区域定义了共形变换. 而实际上, 反演有奇点 $x=x_0$, 在此映射下 (相对于欧氏度量) 它变为 ∞. 而在伪欧氏情形, 由方程 (2) 可明显看出奇点集合为

$$\langle x-x_0,x-x_0\rangle=0,$$

即通过点 x_0 的光锥.

我们还看到, 对于所有相差了一个因子 $\lambda(x)$ 的度量 $g_{\alpha\beta}$ 共形变换群是相同的. 因此我们可以在具黎曼度量的任意空间上对共形变换进行法化, 使它共形等价于在欧氏坐标下的欧氏度量 $g_{\alpha\beta}=\lambda(x)\delta_{\alpha\beta}$. 特别, 由方程

$$\sum_{a=0}^n(x^a)^2=1$$

给出的 n 维球面 S^n 的度量共形等价于在其他某个坐标下的欧氏度量. 当 $n=2$ 时, 在 §9 中已经给出了这样的坐标. 在共形等价于球面 (除去上面的极点) 的任意维的空间中, 像 $n=2$ 的情形那样, 引进到平面 $(x^1,\cdots,x^n)$ 上的球极投影, 并以 $x^1,\cdots,x^n$ 表示. 这时的度量有形式 (参看 §9 中 $n=2$ 的情形) 为

$$\begin{aligned} &g_{\alpha\beta}=g(x)\delta_{\alpha\beta}, \quad g(x)=\frac{R^4}{(R^2+r^2)^2}, \\ &r^2=(x^1)^2+\cdots+(x^n)^2, \quad R=\text{常数}. \end{aligned} \tag{21}$$

在这个坐标 $(x^1,\cdots,x^n)$ 下我们可以考虑刘维尔定理中得到的所有共形变换: 这些是群 $O(n)$ 中的变换, 平移 $\mathbb{R}^n\to\mathbb{R}^n$, 伸缩和反演, 还有它们的复合.

另一方面, 在球面 S^n 上存在它的运动群 $O(n+1)$. 我们进一步将球面看作球 $|x|<R$ 的边缘 $S^n=\partial D^{n+1}$, 在其上的坐标 $(x^0,\cdots,x^n)$ 下定义了罗巴切夫斯基度量 (见 §10)

$$\tilde{g}_{ab}=\frac{R^4}{(R^2-r^2)^2}\delta_{ab}, \quad r^2=\sum_{i=0}^{n}(x^i)^2<R^2. \tag{22}$$

在球 D^{n+1} 上罗巴切夫斯基度量的变换群为 $O(1,n+1)$, 这个群比群 $O(n+1)$ 更大:$O(n+1)\subset O(1,n+1)$. 罗巴切夫斯基空间可以实现为闵可夫斯基空间 $\mathbb{R}^{n+2}_1$ 中的曲面, 在此空间中的坐标为 $(z^0,\cdots,z^{n+1})$, 度量为

$$dl^2=(dz^0)^2-\sum_{a=1}^{n+1}(dz^a)^2. \tag{23}$$

罗巴切夫斯基空间 L^{n+1} 以方程

$$(z^0)^2-\sum_{a=1}^{n+1}(z^a)^2=1 \tag{24}$$

表出. 在坐标 $(z^1,\cdots,z^{n+1})$ 与坐标 $(x^0,\cdots,x^n)$ 之间的关联 (参看 (10.6), 其中 $x^0=u, x^1=v, z^1=x, z^2=y$) 为

$$z^0=\frac{2R^2x^{a-1}}{R^2-r^2}, \quad a=1,\cdots,n+1, \quad r^2=\sum_{a=1}^{n+1}(x^{a-1})^2. \tag{25}$$

于是成立下面的论断:

定理 15.2. 洛伦兹群 $O(1,n+1)$ 在罗巴切夫斯基空间 L^{n+1} 上的作用产生了在 S^n 上的作用 (例如, 如果 L^{n+1} 表现为边界是 S^n 的球 D^{n+1}, 并有度量 (22)). 所有这些变换在球面 S^n 上的标准度量 (21) 下都为共形, 并且没有奇点, 还包含了那些基本变换. 这表明,S^n(或 $\mathbb{R}^n$) 的所有共形变换群因而同构于 $O(1,n+1)$.

为此必须比较 $\mathbb{R}^n$ 的平移, 伸缩和反演与群 $O(1,n+1)$ 或 $O(n+1,1)$ 中的变换. 所有这些变换在 $n=1$ 时已有表示, 它在计算上是极其简单的:

$$\text{平移}(x\mapsto x+a)\leftrightarrow\begin{pmatrix}\dfrac{a^2}{2}+1 & a & -\dfrac{a^2}{2}\\ a & 1 & -a\\ \dfrac{a^2}{2} & a & 1-\dfrac{a^2}{2}\end{pmatrix}\in O(2,1),$$

$$\text{伸缩}(x\mapsto \pm\lambda x,\lambda>0)\leftrightarrow\begin{pmatrix}\dfrac{\lambda^2+1}{2\lambda} & 0 & \dfrac{\lambda^2-1}{2\lambda}\\ 0 & \pm1 & 0\\ \dfrac{\lambda^2-1}{2\lambda} & 0 & \dfrac{\lambda^2+1}{2\lambda}\end{pmatrix}\in O(2,1),$$

$$\text{反演}(x\mapsto \frac{1}{x})\leftrightarrow\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -1\end{pmatrix}\in O(2,1).$$

这些变换在球面 S^1 上均共形, 并生成了整个群 $O(2,1)$.

习题 15.2. 以类似的方式证明空间 $\mathbb{R}^n_1$ 的共形变换群同构于群 $O(n,2)$. 一般地, 空间 $\mathbb{R}^n_{p,q}$ 的共形变换构成了群 $O(p+1,q+1)$. 特别, 对于闵可夫斯基空间 $\mathbb{R}^4_1$, 其共形变换群同构于 $O(4,2)$.

注 另外, 可以证明群 $O(4,2)$ 局部同构于群 $SU(2,2)$. 我们不再构造这个同构了.

第三章　张量. 代数理论

§16. 张量的例子

我们已经习惯于把许多量以空间中点的数值函数来表示, 例如一个点到某个固定中心的距离等等. 如果我们有了某些这样的量, 那么我们便有了某个点的函数了(或者说是点的向量函数). 在三维空间中, 为了完全描述出空间中点的位置至少需要知道三个数值函数, 即称为点的坐标 (x^1, x^2, x^3); 每个坐标 x^i 是点的函数, 而集合 (x^1, x^2, x^3) 则完全确定了这个点. 我们遇到过各种类型的坐标, 譬如在平面上的笛卡儿坐标 x^1, x^2, 极坐标 r, φ, 其中 $x^1 = r\cos\ \varphi, x^2 = r\sin\ \varphi$; 在空间中有笛卡儿坐标, 柱面坐标 r, z, φ 以及球面坐标 r, θ, φ.

于是, 坐标是点的这样一个数值函数组, 它完全确定了此点在空间中的位置. 同样地, 对于一个物理系统, 我们称完全确定此系统的状态的这种数值函数组为它的坐标 (系统的状态即系统 "所有可能状态的空间" 中的点). 例如运动的质点的坐标由 6 个数决定: 3 个坐标及速度向量的 3 个分量; 在这里我们便有了状态的 6 维空间.

然而我们发现, 点的数值函数或者这些函数的集合的概念仍是不充分的. 问题在于许多几何和物理量只在空间中已经给出了某个坐标组 (x^1, x^2, x^3) 之后才能以数值函数组的形式来描述; 如果我们有了另外的坐标 z^1, z^2, z^3,

$$x^i = x^i(z^1, z^2, z^3), \quad i = 1, 2, 3,$$

那么这些量的数值描述可能会有很大的变化. 为了解释清楚这种可能性, 让我们考虑向量的概念, 例如, 沿着某条曲线运动的速度向量, 曲线为

$$z^j = z^j(t), \quad j = 1, 2, 3.$$

在坐标 z^1, z^2, z^3 下速度向量的分量为

$$\left(\frac{dz^1}{dt}, \frac{dz^2}{dt}, \frac{dz^3}{dt}\right)\bigg|_{t=t_0} = (\eta^1, \eta^2, \eta^3).$$

在另一组坐标 x^1, x^2, x^3 下这同一条曲线可写为

$$x^i = x^i(z^1(t), z^2(t), z^3(t)) = x^i(t), \quad i = 1, 2, 3.$$

我们得到这同一向量的另一组分量

$$\left(\frac{dx^1}{dt}, \frac{dx^2}{dt}, \frac{dx^3}{dt}\right)\bigg|_{t=0} = (\xi^1, \xi^2, \xi^3),$$

其中

$$\frac{dx^i}{dt} = \sum_{j=1}^{3} \frac{\partial x^i}{\partial z^j}\frac{dz^j}{dt}, \quad i = 1, 2, 3.$$

于是对此向量的分量我们在坐标变换下有其数值形式的变换公式

$$\begin{aligned} \xi^i &= \sum_{j=1}^{3} \eta^j \frac{\partial x^i}{\partial z^j}, \quad i = 1, 2, 3; \\ x^i &= x^i(z^1, z^2, z^3); \end{aligned} \tag{1}$$

这里的 ξ^1, ξ^2, ξ^3 为此向量在所给点上以坐标 x^1, x^2, x^3 表示的分量, 而 η^1, η^2, η^3 为同一点在坐标 z^1, z^2, z^3 下此向量的分量.

在坐标变换下改变的数值表达式的数量类中, 张量是最重要的一个. 向量是最简单和最直观的张量的例子. 当然, 标量是张量的平凡的例子, 它在坐标变换下不变.

在引进张量的准确的数学概念前, 我们考虑另外一些曾多次遇见过的那些例子.

1. 数值函数的梯度

正像我们已习惯于谈及和想到的那样, 在笛卡儿坐标 x^1, x^2, x^3 下数值函数 $f(x^1, x^2, x^3)$ 的梯度是个向量, 其分量为

$$\text{grad } f = \left(\frac{\partial f}{\partial x^1}, \frac{\partial f}{\partial x^2}, \frac{\partial f}{\partial x^3}\right) = (\xi^1, \xi^2, \xi^3).$$

让我们看一看, 在另外坐标 z^1, z^2, z^3 下同一函数的张量是如何表现的, 其中

$$x^i = x^i(z^1, z^2, z^3) = x^i(z), \quad i = 1, 2, 3.$$

我们有

$$\text{grad } f(x^1(z), x^2(z), x^3(z)) = \left(\frac{\partial f}{\partial z^1}, \frac{\partial f}{\partial z^2}, \frac{\partial f}{\partial z^3}\right) = (\eta_1, \eta_2, \eta_3);$$

$$\frac{\partial f}{\partial z^i}=\sum_{j=1}^{3}\frac{\partial f}{\partial x^j}\frac{\partial x^j}{\partial z^i},\quad i=1,2,3.$$

因而我们的答案是

$$\eta_i=\sum_{j=1}^{3}\frac{\partial x^j}{\partial z^i}\xi_j;\tag{2}$$

这里的 ξ_1,ξ_2,ξ_3 为在坐标 x^1,x^2,x^3 下这个函数的梯度分量,η_1,η_2,η_3 则是在坐标 z^1,z^2,z^3 下的梯度分量.

现在比较曲线的速度向量和函数的梯度的数值表达式的变换公式.

$$\text{速度向量:}\quad \xi^i=\sum_{j=1}^{3}\eta^j\frac{\partial x^i}{\partial z^j}.$$

$$\text{梯度:}\quad \eta_i=\sum_{j=1}^{3}\xi_j\frac{\partial x^j}{\partial z^i}.$$

这是些不同的公式!

为了比较这两个公式, 我们引进雅可比矩阵 $A=(a_j^i)$, 其中 $a_j^i=\dfrac{\partial x^i}{\partial z^j}$ 和转置矩阵 $A^{\mathrm T}=(b_k^j)$, 其中 $b_k^j=a_j^k$.

对于向量 ξ 和 η 我们把公式 (1) 和 (2) 重写为形式

$$\xi=A\eta\ (\text{速度向量}),\tag{1'}$$

$$\eta=A^{\mathrm T}\xi\ (\text{梯度}).\tag{2'}$$

如果矩阵 $A^{\mathrm T}$ 有逆 $(A^{\mathrm T})^{-1}$, 我们则可重写公式 (2),(2′) 为

$$\begin{aligned}&(A^{\mathrm T})^{-1}\eta=(A^{\mathrm T})^{-1}A^{\mathrm T}\xi=\xi,\\&\xi=(A^{\mathrm T})^{-1}\eta\ \left(\text{或 }\xi_i=\sum_{j=1}^{3}\eta_j\frac{\partial z^j}{\partial x^i}\right).\end{aligned}\tag{3}$$

在什么情形下速度向量和函数的梯度当由坐标系 x 变到坐标 z 时的变换规律相同?

由公式 (1),(2) 和 (3) 我们得到

$$\xi=A\eta\ (\text{速度向量}),$$

$$\xi=(A^{\mathrm T})^{-1}\eta\ (\text{函数的梯度}).$$

最后的结论是, 为了使变换 (1) 和 (3) 相重合必须使矩阵相等:

$$A=(A^{\mathrm T})^{-1}\quad\text{或者}\quad A^{\mathrm T}A=1,$$

即 $A=(a^i_j)=\left(\frac{\partial x^i}{\partial z^j}\right)$ 为正交矩阵 (参看 §4).(附带给个评注: 使得在每点的雅可比矩阵 $A=\left(\frac{\partial x^i}{\partial z^j}\right)$ 为正交的坐标变换 $x=x(z)$ 是个线性的正交变换:$A=$ 常数矩阵.)

因此, 在坐标变换下函数的梯度与速度向量的变化不同. 这是另一种向量形式, 我们称之为 "余向量" 以区别于速度向量.

2. 黎曼度量

像已经说过的那样, 在给出坐标 $(x^1,\cdots,x^n)$ 的 n 维空间中度量概念 (诸如长度, 角度) 借助于函数组 $g_{ij}(x),i=1,\cdots,n,j=1,\cdots,n$ 给出. 按定义, 曲线的长度由公式

$$l=\int_a^b\sqrt{\sum_{i,j}g_{ij}(x(t))\dot{x}^i\dot{x}^j}dt \tag{4}$$

定义, 其中 $\dot{x}=\frac{dx}{dt}$, 且二次型 $\sum g_{ij}\xi^i\xi^j$ 正定. 这是一个定义于在每个所给点 $(x)=(x^1,\cdots,x^n)$ 上的 "速度向量" 型向量上的二次型, 它依赖于点. 我们称 g_{ij} 为黎曼度量. 在坐标变换

$$x^i=x^i(z^1,\cdots,z^n),\quad i=1,\cdots,n,$$

下的曲线的长度公式有了形式

$$l=\int_a^b\sqrt{\sum_{i,j}\tilde{g}_{ij}(z(t))\dot{z}^i\dot{z}^j}\;dt,$$

其中 $x^i(t)=x^i(z^1(t),\cdots,z^n(t))$, 同时度量分量的变换规律的形式为

$$\tilde{g}_{ij}(z)=\sum_{k,l=1}^n g_{kl}(x)\frac{\partial x^k}{\partial z^i}\frac{\partial x^l}{\partial z^j}. \tag{5}$$

因此, 向量的二次型按规律 (5) 变化. 这也是张量的一种形式 (称之为 2 阶张量).

于是, 我们已经有了几种类型的张量:

a) 标量 (不变化),

b) 向量 (按规则 (1) 变化),

c) 余向量 (按规则 (2) 变化),

d) 黎曼度量 (按规则 (5) 变化).

我们记起在所给坐标 $(x^1,\cdots,x^n)$ 下黎曼度量对于定义在所给点上的长度概念是不可少的: 如果在点 $x=(x^1,\cdots,x^n)$ 给出向量 $\xi=(\xi^1,\cdots,\xi^n)$, 则长度的平方等于 $\sum g_{ij}(x)\xi^i\xi^j$. 特别, 它用于参数化的曲线的速度向量上, 以将曲线的长度定义为速度向量长度的积分.

对于在坐标变换下度量分量的变化规则 (5) 能唯一地从向量分量的规则 (1) 和明显的一种要求得出, 这个要求是要使得曲线的长度不依赖于进行这个计算的坐标. 曲线的真正长度是速度向量长度按时间的积分. 因此必须使速度向量长度平方

$$|\xi|^2=\sum g_{ij}\xi^i\xi^j$$

不依赖于坐标的选取. 由此要求以及由对向量的公式 (1), 我们得到对度量 (g_{ij}) 的分量的变化规律 (5).

3. 如果我们希望定义出按规律 (2)(或 (3)) 变化的余向量长度平方的不变概念, 我们则应该引进分量 $(g^{ij}_{(x)})$ 并令

$$|\xi|^2=\sum g^{ij}(x)\xi_i\xi_j\ (\text{于点}x),\quad \xi=(\xi_1,\cdots,\xi_n).$$

在变换 $x^i=x^i(z), i=1,\cdots,n, \eta_i=\sum\limits_{j=1}^n\xi_j\dfrac{\partial x^j}{\partial z^i}$ 下, 我们得到下面的变换规则

$$\bar g^{ij}(z)=\sum_{k,l=1}^n g^{kj}\frac{\partial z^i}{\partial x^k}\frac{\partial z^j}{\partial x^l};\tag{6}$$

长度将不依赖于坐标的选取

$$|\eta|^2=|\xi|^2=\sum\bar g^{ij}\eta_i\eta_j=\sum g^{ij}\xi_i\xi_j.$$

这里 $\xi=(\xi_1,\cdots,\xi_n)$ 是在坐标 $x^1,\cdots,x^n$ 下在点 (x) 的余向量的记号.$\eta=(\eta_1,\cdots,\eta_n)$ 为在同一点上同一余向量的记号, 但是是在坐标 $z^1,\cdots,z^n$ 下的.

变换规律 (6) 又给出了一种张量类型 (2 阶).

4. 最后应该分析一下仍然缺失的一个二阶张量类型, 即向量的线性算子.

假设在具坐标 $x^1,\cdots,x^n$ 的空间中的每个点上给出了矩阵 $(a^i_j(x))=A(x)$, 它在每点 $x=(x^1,\cdots,x^n)$ 定义了向量的线性变换. 这个线性变换 $A(x)$ 有形式 $\eta=A\xi$, 其中

$$\eta^i=\sum_{j=1}^n a^i_j(x)\xi^j,\tag{7}$$

其中 $\xi=(\xi^1,\cdots,\xi^n)$ 为在点 x 的向量.

这个同一的矩阵定义了余向量的线性变换, 其公式为 $\eta=A\xi$, 其中

$$\eta_j=\sum_{i=1}^n a^i_j\xi_i.\tag{8}$$

在坐标变换 $x^i=x^i(z^1,\cdots,z^n)$ 下由公式 (1),(7) 可得出矩阵 A 的分量按 $A\mapsto\tilde A=(\tilde a^i_j)$ 的规则变化:

$$\tilde a^i_j=\sum\frac{\partial z^i}{\partial x^k}a^k_l\frac{\partial x^l}{\partial z^j},\tag{9}$$

其中 $x^i = x^i(z), z^j = z^j(x)$ 和 $z^i(x(z)) = z^i$; 于是

$$\sum_{k=1}^{n} \frac{\partial z^i}{\partial x^k} \frac{\partial x^k}{\partial z^j} = \delta^i_j = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases}$$

对于余向量我们可将公式 (2) 重写为

$$\xi_i = \sum \eta_j \frac{\partial z^j}{\partial x^i},$$

这是因为有

$$\sum_{i=1}^{n} \frac{\partial z^j}{\partial x^i} \frac{\partial x^i}{\partial z^k} = \delta^j_k.$$

现在把向量, 余向量和所有上述三类 2 阶张量形式的变换规则集中列于下面.

变化规则:

1. 标量 (0 阶张量); 不变化.

1 阶张量:

2. 向量 $\xi = (\xi^i)$ (速度向量型); 变换公式为

$$\tilde{\xi}^j = \sum_{i=1}^{n} \frac{\partial z^j}{\partial x^i} \xi^i.$$

3. 余向量 $\xi = (\xi_i)$(函数梯度型); 变换公式为

$$\tilde{\xi}_j = \sum_{i=1}^{n} \xi_i \frac{\partial x^i}{\partial z^j}.$$

2 阶张量:

4. 向量的内积 g_{ij}; 变换公式为

$$\tilde{g}_{ij} = \sum_{k,l} g_{kl} \frac{\partial x^k}{\partial z^i} \frac{\partial x^l}{\partial z^j}.$$

5. 余向量的内积 g^{ij}; 变换公式为

$$\tilde{g}^{ij} = \sum_{k,l} g^{kl} \frac{\partial z^i}{\partial x^k} \frac{\partial z^j}{\partial x^l}.$$

6. 向量 (余向量) 的线性算子$A = (a^i_j)$; 变换公式为

$$\tilde{a}^i_j = \sum_{k,l} a^k_l \frac{\partial x^l}{\partial z^j} \frac{\partial z^i}{\partial x^k}.$$

在这里 $x^i = x^i(z^1, \cdots, z^n), z^j = z^j(x^1, \cdots, x^n)$, 另外

$$z^j(x^1(z^1, \cdots, z^n), \cdots, x^n(z^1, \cdots, z^n)) = z^j;$$

$$\sum_{k=1}^{n} \frac{\partial z^i}{\partial x^k} \frac{\partial x^k}{\partial z^j} = \delta^i_j.$$

§17. 张量的一般定义

1. 任意阶张量的分量的变换规律

在上一节中我们分析了一阶张量 (向量和余向量) 和 2 阶张量 (向量的二次型 g_{ij}, 余向量的二次型 g^{ij} 和线性变换即算子 a^i_j).

现在我们可以给出具任意阶的张量的定义.

定义 17.1. 我们称对任意一个坐标系 $(x^1,\cdots,x^n)$ 给出的数组 $T^{i_1\cdots i_p}_{j_1\cdots j_q}$ 为 $p+q$阶的(p,q)型张量 (张量场)是说这些数值记号按下面的规则依赖于坐标系: 如果 $x^i=x^i(z^1,\cdots,z^n), z^j=z^j(x^1,\cdots,x^n); z(x(z))=z$, 则有公式

$$T^{i_1\cdots i_p}_{j_1\cdots j_q}=\sum_{(k),(l)}{}'T^{k_1\cdots k_p}_{l_1\cdots l_q}\frac{\partial x^{i_1}}{\partial z^{k_1}}\cdots\frac{\partial x^{i_p}}{\partial z^{k_p}}\frac{\partial z^{l_1}}{\partial x^{j_1}}\cdots\frac{\partial z^{l_q}}{\partial x^{j_q}};\qquad(1)$$

这里的 ${}'T^{k_1\cdots k_p}_{l_1\cdots l_q}$ 是此张量在坐标 (z) 下的数值记号 (分量), 而 $T^{i_1\cdots i_p}_{j_1\cdots j_q}$ 为在坐标 (x) 下的数值记号.

指标 $(i_1,\cdots,i_p;j_1,\cdots,j_q)$ 和 $(k_1,\cdots,k_p;l_1,\cdots,l_q)$ 在 1 到 n 间变化 (这是在 n 维空间中的张量).

因此:

速度向量是 (1,0) 型张量,

余向量是 (0,1) 型张量,

向量的二次型是 (0,2) 型张量,

余向量的二次型是 (2,0) 型张量,

向量或余向量的线性算子是 (1,1) 型张量.

定理 17.1. 分量 ${}'T^{k_1\cdots k_p}_{l_1\cdots l_q}$ 可以通过 $T^{i_1\cdots i_p}_{j_1\cdots j_q}$ 由下面的公式表出

$${}'T^{k_1\cdots k_p}_{l_1\cdots l_q}=\sum_{(i),(j)}T^{i_1\cdots i_p}_{j_1\cdots j_q}\frac{\partial z^{k_1}}{\partial x^{i_1}}\cdots\frac{\partial z^{k_p}}{\partial x^{i_p}}\frac{\partial x^{j_1}}{\partial z^{l_1}}\cdots\frac{\partial x^{j_q}}{\partial z^{l_q}}.\qquad(2)$$

证明

我们要利用下面的关系式

$$\sum_j\frac{\partial x^i}{\partial z^j}\frac{\partial z^j}{\partial x^k}=\delta^i_k,\quad \sum_j\frac{\partial z^k}{\partial x^j}\frac{\partial x^j}{\partial z^q}=\delta^k_q,$$

它是由变换 $x=x(z)$ 和 $z=z(x)$ 互逆得到, 即

$$x^i(z(x))=x^i;\quad z^q(x(z))=z^q.$$

考虑关系式 (1) 为右端是 $T^{i_1\cdots i_p}_{j_1\cdots j_q}$ 而未知量是 ${}'T^{k_1\cdots k_p}_{l_1\cdots l_q}$ 的线性方程. 解此方程必会得到 (2).

由 (1) 成立有公式

$$\begin{aligned}&\sum_{ij} T_j^i \frac{\partial z^{k_1}}{\partial x^{i_1}} \cdots \frac{\partial z^{k_p}}{\partial x^{i_p}} \frac{\partial x^{j_1}}{\partial z^{l_1}} \cdots \frac{\partial x^{j_q}}{\partial z^{l_q}}\\ =&\sum_{i,j}\left(\sum_{r,s} {}'T_s^r \frac{\partial x^{i_1}}{\partial z^{r_1}} \cdots \frac{\partial x^{i_p}}{\partial z^{r_p}} \frac{\partial z^{s_1}}{\partial x^{j_1}} \cdots \frac{\partial z^{s_q}}{\partial x^{j_q}}\right) \frac{\partial z^{k_1}}{\partial x^{i_1}} \cdots \frac{\partial z^{k_p}}{\partial x^{i_p}} \frac{\partial x^{j_1}}{\partial z^{l_1}} \cdots \frac{\partial x^{j_q}}{\partial z^{l_q}}\\ =&\sum_{i,j,r,s} {}'T_s^r \frac{\partial x^i}{\partial z^r} \frac{\partial z^s}{\partial x^j} \frac{\partial z^k}{\partial x^i} \frac{\partial x^j}{\partial z^l} = \sum_{r,s} {}'T_s^r \delta_l^s \delta_r^k =' T_l^k;\end{aligned}$$

其中

$$i = (i_1, \cdots, i_p); \quad j = (j_1, \cdots, j_q);$$

$$l = (l_1, \cdots, l_q); \quad s = (s_1, \cdots, s_q);$$

$$k = (k_1, \cdots, k_p); \quad r = (r_1, \cdots, r_p).$$

故而变成了关系 (2). 定理证完. □

我们列出张量的最简单的性质.

在空间的任意给定点上张量构成了线性空间: 如果 $T = (T_{j_1\cdots j_q}^{i_1\cdots i_p}), S = (S_{j_1\cdots j_q}^{i_1\cdots i_p})$ 为 (p,q) 型张量, 于是它们的线性组合 $\lambda T + \mu S = U$ 在同一点上也是 (p,q) 型的张量, 其中分量 $U_{j_1\cdots j_q}^{i_1\cdots i_p} = \lambda T_{j_1\cdots j_q}^{i_1\cdots i_p} + \mu S_{j_1\cdots j_q}^{i_1\cdots i_p}$. 必须指出, 张量是从属于点的一个对象, 而不存在从属于不同点的张量的加法规律.

在 n 维空间中的 (p,q) 型 (在一点上的) 张量的线性空间, 其维数等于 n^{p+q}. 如果在坐标系为 $(x^1, \cdots, x^n)$ 的 n 维空间中基坐标向量以 $e_1, \cdots, e_n$ 表示, 而基余向量为 $e^1, \cdots, e^n$, 则任意张量可方便地从形式上写为下列形状:

向量 $\xi = \sum_i \xi^i e_i \left(例如, \dfrac{dx}{dt} = \sum \dfrac{dx^i}{dt} e_i\right)$,

余向量 $\xi = \sum_i \xi_i e^i \left(例如, \mathrm{grad}\, f = \sum \dfrac{\partial f}{\partial x^i} e^i\right)$,

二次型 $(g_{ij}) = \sum_{i,j} g_{ij} e^i \otimes e^j$ (对向量),

二次型 $(g^{ij}) = \sum_{i,j} g^{ij} e_i \otimes e_j$(对余向量),

线性算子 $A = \sum_{i,j} a_j^i e_i \otimes e^j$.

任一个张量 $T = (T_{j_1\cdots j_q}^{i_1\cdots i_p})$ 可以写成

$$T = \sum_{i,j} T_{j_1\cdots j_q}^{i_1\cdots i_p} e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes e^{j_1} \otimes \cdots \otimes e^{j_q}. \tag{3}$$

值得注意的是, 在这些记号中指标的顺序是重要的: 不能交换位置, 譬如说, e_{i_1} 和 e_{i_2} 等等.

因此, 在空间中点 (x) 的 (p,q) 型张量的线性空间中的基具有形状

$$e_{i_1}\otimes\cdots\otimes e_{i_p}\otimes e^{j_1}\otimes\cdots\otimes e^{j_q}, \tag{4}$$

其中 i, j 独立地取值于 $1,\cdots,n$; 因而基由 n^{p+q} 个元构成. 在坐标变换 $x^i = x^i(z^1,\cdots,z^n)$ 下我们过渡到从属于同一点的张量的线性空间的另一组基, 即在所给点上与坐标向量系 z 相关联的基.

在空间中所给点上这些由一个到另一个基的相互表达式按下面公式进行:

$$e_{i_1}\otimes\cdots\otimes e_{i_p}\otimes e^{j_1}\otimes\cdots\otimes e^{j_q}$$

$$= \sum_{(k),(l)}\frac{\partial z^{k_1'}}{\partial x^{i_1}}\cdots\frac{\partial z^{k_p'}}{\partial x^{i_p}}\frac{\partial x^{j_1}}{\partial z^{l_1'}}\cdots\frac{\partial x^{j_q}}{\partial z^{l_q'}}e_{k_1'}\otimes\cdots\otimes e_{k_p'}\otimes e^{l_1'}\otimes\cdots\otimes e^{l_q'}. \tag{4'}$$

我们列举几个例子.

1) *应力张量 (三维情形)*. 在连续介质的每点 $x=(x^1,x^2,x^3)$, 作用于介质的小面积 ΔS 上的张力由 $(\Delta S)P(n)$ 计算, 其中 ΔS 垂直于单位向量 n, P 为线性算子: $P=(P_j^i)$. 称张量 P_j^i 为*应力张量*. 如果 $n=\sum\limits_{j=1}^{3}n^je_j$, 则 $P(n)=\sum\limits_{i=1}^{3}\left(\sum\limits_{j=1}^{3}n^jP_j^i\right)e_i$ 或 $\{P(n)\}^i=\sum\limits_{j=1}^{3}n^jP_j^i$. 例如, 在可应用帕斯卡定律的场合, 即作用于单位正交面积上的力与方向无关, $P_j^i=\delta_j^ip$, 其中 p 是被称为所给点上的张力的数值.

2) *形变张量*. 如果连续介质在坐标 x^1,x^2,x^3 中给出, 并在介质的每点上给出了一个位移

$$x^i\longmapsto x^i+u^i(x),$$

就是说, 介质经受了形变. 如果在初始状态下介质的邻近点之间的距离, 譬如, 在坐标 x^1,x^2,x^3 下为欧氏的

$$(\Delta l)^2=\sum(\Delta x^i)^2=\sum_{i=1}^{3}(x^i-{}'x^i)^2,$$

则在形变之后的在同样的两点之间有另一种距离 $(\Delta\bar{l})^2=\sum[x^i+u^i(x)-'x^i-u^i('x)]^2$. 显然

$$(\Delta\bar{l})^2=(\Delta l)^2+2\sum_{i=1}^{3}\Delta x^i\Delta u^i+\sum_{i=1}^{3}(\Delta u^i)^2,$$
$$\Delta x^i=x^i-'x^i,$$
$$\Delta u^i=u^i(x)-u^i('x)\approx\sum_{j=1}^{3}\frac{\partial u^i}{\partial x^j}\Delta x^j.$$

于是当 $\Delta x^i\to 0$ 时

$$(d\bar{l})^2=(dl)^2+2\sum_{i,j}dx^idx^j\frac{\partial u^i}{\partial x^j}+\sum_{k,l,i}\frac{\partial u^i}{\partial x^k}\frac{\partial u^i}{\partial x^l}dx^kdx^l.$$

因为

$$\sum_{i,j}\frac{\partial u^i}{\partial x^j}dx^idx^j=\sum_{i,j}\frac{\partial u^j}{\partial x^i}dx^idx^j,$$

则有等式

$$2\sum_{i,j}\frac{\partial u^i}{\partial x^j}dx^idx^j=\sum_{i,j}\left(\frac{\partial u^i}{\partial x^j}+\frac{\partial u^j}{\partial x^i}\right)dx^idx^j.$$

于是

$$(d\bar{l})^2-(dl)^2=\sum_{i,j}\left(\eta_{ij}+\sum_k\frac{\partial u^k}{\partial x^i}\frac{\partial u^k}{\partial x^j}\right)dx^idx^j.$$

定义 17.2. 称系数 $\eta_{ij}+\sum\limits_k\frac{\partial u^k}{\partial x^i}\frac{\partial u^k}{\partial x^j}$ 为介质的形变张量, 其中 $\eta_{ij}=\frac{\partial u^i}{\partial x^j}+\frac{\partial u^j}{\partial x^i}$.

如果 u^i 为小位移, 则关于 u^i 的二次部分可以忽略不计. 我们因而得到了小形变张量

$$\eta_{ij}=\frac{\partial u^i}{\partial x^j}+\frac{\partial u^j}{\partial x^i}.$$

根据胡克定律, 小的介质形变所产生的应力线性依赖于形变. 所以在应力张量和形变之间应是线性关联的

$$P=U(\eta).$$

这个线性关联是个 4 阶张量. 在指标形式下有 $P^i_j=\sum\limits_{k,l}U^{ikl}_j\eta_{kl}$; 这里的 $P=(P^i_j),\eta=(\eta_{kl}),U=(U^{ikl}_j)$. 张量 U^{ikl}_j 为 4 阶, 由 81 个分量描述. 在连续介质上的胡克定律难道真的要 81 个数量来描述吗?

如果坐标是欧几里得的, 我们则可以对向量和余向量 (相对于正交变换) 不加区别, 或更一般地对张量的上指标和下指标不加区别, 它们的变换是相同的. 张量

$U = (U_j^{ikl})$ 在欧氏坐标下定义了 81 个数. 在连续介质的各向同性 (迷向) 的假设下, 这个数目可以大大缩减. 这表明在每点, 张量 U 在只相差一个绕此点的旋转 (包括反射) 的所有坐标下, 它的这种数值记号是唯一的, 也就是说, 在正交变换下不变. 这个假设在流体情形成立, 然而对刚体则不总成立.

在各向同性成立的条件下可以利用下面的这个重要定理, 然而我们只叙述于后而不加以证明 (由于问题是关于正交变换的, 故而我们对于上下指标不加以区别).

定理 17.2. 在正交变换下不变的 4 阶张量类由三个参数 λ, μ, ν 决定; 它由张量

$$U_{ijkl} = \lambda\delta_{ik}\delta_{jl} + \mu\delta_{ij}\delta_{kl} + \nu\delta_{il}\delta_{jk}$$

构成.

用从双指标张量 η_{ij} 变到双指标张量 P_{ij} 的方式表示, 则这个张量可由公式

$$P_{ij} = \lambda\eta_{ij} + \mu(\mathrm{Tr}\ \eta)\delta_{ij} + \nu\eta_{ji}$$

描述, 其中 $\mathrm{Tr}\ \eta = \sum\limits_i \eta_{ii}$. 如果我们记起所考虑的张量 η 和 P 是对称的, 即 $\eta_{ij} = \eta_{ji}, P_{ij} = P_{ji}$, 那么张量 P 实际上只依赖于两个参数, 即 $\lambda + \nu$ 和 μ. 最后结论: 在各向同性物质中, 描述 4 阶张量的两个 2 阶的物理的 (对称) 张量之间相关联的线性规则在介质的每个点上只依赖于两个常数.

自然的问题是,1,2 和 3 阶的迷向张量会是什么样.

当然, 不存在那样的向量和余向量, 它的坐标在旋转下不变.

至于 2 阶张量, 在具坐标 x, y 的二维空间的任意张量 (h_{ij}), 在变换 $x \mapsto x, y \to -y$ 下变为 $(h'_{ij}) : h'_{11} = h_{11}, h'_{22} = h_{22}, h'_{12} = -h_{12}, h'_{21} = -h_{21}$, 而在变换 $x \mapsto y, y \mapsto x$ 下变为张量 $(h''_{ij}), h''_{11} = h_{22}, h''_{22} = h_{11}, h''_{12} = h_{21}, h''_{21} = h_{12}$; 使 $h_{ij} = h'_{ij} = h''_{ij}$ 给出 $h_{11} = h_{22}, h_{12} = h_{21} = 0$, 即 $h_{ij} = \lambda\delta_{ij}$. 因此,$(\lambda\delta_{ij})$ 是唯一的 2 阶迷向张量 (显然, 提到的上述论证容易推广到任意维空间的情形).

可以证明不存在 3 阶的迷向张量.

我们要强调指出, 迷向性的假设要假定黎曼度量的存在 (我们明确假定的是欧氏度量), 然而张量的概念与度量并无关联; 度量本身就是一种张量. 所以值得看一看, 哪些张量不仅在旋转下不变, 而且更一般地对于线性变换不变. 这时区分上和下指标就是重要的了.

首先明白的是,(p,q) 阶的非零张量只有在 $p = q$ 时才能是不变的: 当坐标系在相似伸缩 $\lambda \neq 1$ 时张量的每个分量均乘以 λ^{q-p}, 它只有在 $p - q = 0$ 时才能等于 1. 特别, 不变张量必具偶数阶. 在 2 阶张量中不变的只有张量 $\eta_j^i = \lambda\delta_j^i$: 已是旋转下不变的 2 阶张量不会再多了. 结果, 不变的 4 阶张量是具有两个参数的张量族 $U_{kl}^{ij} = \lambda\delta_k^i\delta_l^j + \mu\delta_l^i\delta_k^j$.

2. 张量的代数运算

我们先引进一些方便的符号. 设在坐标系 $(x^1,\cdots,x^n)$ 下给出了 (p,q) 型张量的分量 $T^{i_1\cdots i_p}_{j_1\cdots j_q}$. 设给出了另一坐标系 $(x^{1'},\cdots,x^{n'})$(其中撇 "′" 在指标上), 其中

$$x^{k'}=x^{k'}(x^1,\cdots,x^n),\quad k'=1,\cdots,n. \tag{5}$$

我们张量的分量在加撇的坐标系中以 $T^{i'_1\cdots i'_p}_{j'_1\cdots j'_q}$ 表示. 这时变换规则 (1) 有形式

$$T^{i'_1\cdots i'_p}_{j'_1\cdots j'_q}=\sum_{(i),(j)}T^{i_1\cdots i_p}_{j_1\cdots j_q}\frac{\partial x^{i'_1}}{\partial x^{i_1}}\cdots\frac{\partial x^{i'_p}}{\partial x^{i_p}}\frac{\partial x^{j_1}}{\partial x^{j'_1}}\cdots\frac{\partial x^{j_q}}{\partial x^{j'_q}} \tag{6}$$

或者

$$T^{i_1\cdots i_p}_{j_1\cdots j_q}=\sum_{(i'),(j')}T^{i'_1\cdots i'_p}_{j'_1\cdots j'_q}\frac{\partial x^{i_1}}{\partial x^{i'_1}}\cdots\frac{\partial x^{i_p}}{\partial x^{i'_p}}\frac{\partial x^{j'_1}}{\partial x^{j_1}}\cdots\frac{\partial x^{j'_q}}{\partial x^{j_q}}. \tag{7}$$

另外, 我们有下面的张量约化规则: 如果在一个公式中指标出现了两重, 则表示按此指标从 1 到 n(n 为空间的维数) 求和, 而求和记号 $\sum$ 则可以不再写出. 那么, 在公式 (6) 中要按两次重复的指标 (上和下面的)$i_1,\cdots,i_p;j_1,\cdots,j_q$ 取和; 在公式 (7) 中二次重复指标为 $i'_1,\cdots,i'_p;j'_1,\cdots,j'_q$; 而指标 i_1 和 i'_1 等等则是独立的.

运用这个规则, 还有加撇的指标, 可以在书写张量分析的公式时避免出错.

现在我们来引进三个重要的张量的代数运算. 我们先在一个固定的坐标系 $(x^1,\cdots,x^n)$ 下定义它们.

1) 指标置换. 设 σ 为数 $1,\cdots,q$ 的一个置换:$\sigma=\begin{pmatrix}1\cdots q\\ \sigma(1)\cdots\sigma(q)\end{pmatrix}$. 置换 σ 作用于数组 $(j_1,\cdots,j_q)$ 的规则为

$$\sigma(j_1,\cdots,j_q)=(j_{\sigma(1)},\cdots,j_{\sigma(q)}). \tag{8}$$

如果成立

$$\widetilde{T}^{i_1\cdots i_p}_{j_1\cdots j_q}=T^{i_1\cdots i_p}_{\sigma(j_1,\cdots,j_q)}, \tag{9}$$

我们则说张量 $\widetilde{T}^{i_1\cdots i_p}_{j_1\cdots j_q}$ 是由张量 $T^{i_1\cdots i_p}_{j_1\cdots j_q}$ 经下指标的置换得到. 上指标置换的定义是类似的. 上下指标之间则不能置换: 这样的运算对于坐标变换不是不变的. 在二阶张量 (T_{ij}) 的情形,i 和 j 的置换转化为此张量的分量构成的矩阵的转置运算.

2) 缩并 (取迹).(p,q) 型张量 $(T^{i_1\cdots i_p}_{j_1\cdots j_q})$ 对于指标 (i_k,j_l) 的缩并是指张量 $(\widetilde{T}^{i_1\cdots i_{p-1}}_{j_1\cdots j_{q-1}})$, 它是 $(p-1,q-1)$ 型的, 由公式

$$\widetilde{T}^{i_1\cdots i_{p-1}}_{j_1\cdots j_{q-1}}=T^{i_1\cdots i_{k-1}ii_k\cdots i_{p-1}}_{j_1\cdots j_{l-1}ij_l\cdots j_{q-1}} \tag{10}$$

(请记住, 按两次出现的 (上和下) 指标进行从 1 到 n 的求和). 例如,(1,1) 型张量 (T^i_j) 的缩并是标量 T^i_i (即线性运算 (T^i_j) 的迹 $\mathrm{Tr}\,T$).

3) (张量的)乘积. 如果给出了 (p,q) 和 (k,l) 型的两个张量 $(T^{i_1\cdots i_p}_{j_1\cdots j_q})$ 和 $(P^{i_1\cdots i_k}_{j_1\cdots j_l})$, 则它们的乘积 $S=T\otimes P$ 为 $(p+k,q+l)$ 型张量, 其分量为

$$S^{i_1\cdots i_{p+k}}_{j_1\cdots j_{q+l}}=T^{i_1\cdots i_p}_{j_1\cdots j_q}P^{i_{p+1}\cdots i_{p+k}}_{j_{q+1}\cdots j_{q+l}} \tag{11}$$

(乘积因子的次序是有本质作用的). 但我们留意到张量积是可结合的运算.

引理 17.1. 作为运算 1), 2), 3) 所得到的结果仍旧是一个张量, 另外应用这些运算的所得结果不依赖于坐标系的选取.

证明 1) 假设 σ 只置换它们自己的指标 k 和 l

$$\sigma=\begin{pmatrix}1\cdots k\cdots l\cdots q\\ 1\cdots l\cdots k\cdots q\end{pmatrix}.$$

于是

$$\widetilde{T}^{i_1\cdots i_p}_{j_1\cdots j_k\cdots j_l\cdots j_q}=T^{i_1\cdots i_p}_{j_1\cdots j_l\cdots j_k\cdots j_q}. \tag{12}$$

当转换到带撇的坐标系时我们有

$$\begin{aligned}\widetilde{T}^{i_1\cdots i_p}_{j_1\cdots j_q}&=T^{i_1\cdots i_p}_{j_1\cdots j_l\cdots j_k\cdots j_q}\\ &=T^{i'_1\cdots i'_p}_{j'_1\cdots j'_q}\frac{\partial x^{i_1}}{\partial x^{i'_1}}\cdots\frac{\partial x^{i_p}}{\partial x^{i'_p}}\frac{\partial x^{j'_1}}{\partial x^{j_1}}\cdots\frac{\partial x^{j'_k}}{\partial x^{j_l}}\cdots\frac{\partial x^{j'_l}}{\partial x^{j_k}}\cdots\frac{\partial x^{j'_q}}{\partial x^{j_q}}.\end{aligned} \tag{13}$$

改变右端的指标记号: 以 j'_l 表示 j'_k, 以 j'_k 表示 j'_l. 这并没有改变整个表达式, 其原因在于我们在按这些指标进行了求和. 这样, 公式 (13) 的右端便有形式

$$\begin{aligned}&T^{i'_1\cdots i'_p}_{j'_1\cdots j'_l\cdots j'_k\cdots j'_q}\frac{\partial x^{i_1}}{\partial x^{i'_1}}\cdots\frac{\partial x^{i_p}}{\partial x^{i'_p}}\frac{\partial x^{j'_1}}{\partial x^{j_1}}\cdots\frac{\partial x^{j'_l}}{\partial x^{j_l}}\cdots\frac{\partial x^{j'_k}}{\partial x^{j_k}}\cdots\frac{\partial x^{j'_q}}{\partial x^{j_q}}.\\ =&\widetilde{T}^{i'_1\cdots i'_p}_{j'_1\cdots j'_q}\frac{\partial x^{i_1}}{\partial x^{i'_1}}\cdots\frac{\partial x^{i_p}}{\partial x^{i'_p}}\frac{\partial x^{j'_1}}{\partial x^{j_1}}\cdots\frac{\partial x^{j'_q}}{\partial x^{j_q}},\end{aligned}$$

即我们已证明了 $\widetilde{T}^{i_1\cdots i_p}_{j_1\cdots j_q}$ 为 (p,q) 型张量的分量.

2) 对于按指标 i_k 和 j_l 缩并的张量我们有 (以 $\widehat{i_k}$ 和 $\widehat{j_l}$ 表示略去了这些指标)

$$\begin{aligned}\widetilde{T}^{i_1\cdots\widehat{i_k}\cdots i_p}_{j_1\cdots\widehat{j_l}\cdots j_q}&=T^{i'_1\cdots i'_p}_{j'_1\cdots j'_q}\frac{\partial x^{i_1}}{\partial x^{i'_1}}\cdots\frac{\partial x^{i_p}}{\partial x^{i'_p}}\frac{\partial x^{j'_1}}{\partial x^{j_1}}\cdots\frac{\partial x^{j'_q}}{\partial x^{j_q}}\bigg|_{i_k=j_l=i}\\ &=T^{i'_1\cdots i'_p}_{j'_1\cdots j'_q}\delta^{j'_l}_{i'_k}\frac{\partial x^{i_1}}{\partial x^{i'_1}}\cdots\frac{\widehat{\partial x^{i_k}}}{\partial x^{i'_k}}\cdots\frac{\partial x^{i_p}}{\partial x^{i'_p}}\frac{\partial x^{j'_1}}{\partial x^{j_1}}\cdots\frac{\widehat{\partial x^{j'_l}}}{\partial x^{j_l}}\cdots\frac{\partial x^{j'_q}}{\partial x^{j_q}}\\ &=\widetilde{T}^{i'_1\cdots\widehat{i'_k}\cdots i'_p}_{j'_1\cdots\widehat{j'_l}\cdots j'_q}\frac{\partial x^{i_1}}{\partial x^{i'_1}}\cdots\frac{\widehat{\partial x^{i_k}}}{\partial x^{i'_k}}\cdots\frac{\partial x^{i_p}}{\partial x^{i'_p}}\frac{\partial x^{j'_1}}{\partial x^{j_1}}\cdots\frac{\widehat{\partial x^{j'_l}}}{\partial x^{j_l}}\cdots\frac{\partial x^{j'_q}}{\partial x^{j_q}},\end{aligned}$$

其中我们利用了等式 $\dfrac{\partial x^i}{\partial x^{i'_k}}\dfrac{\partial x^{j'_l}}{\partial x^i}=\delta^{j'_l}_{i'_k}$. 乘积 (11) 的张量性质是显然的. 引理得证. □

现在考虑所构建的运算的应用例子.

例.

17.1 设已给向量 ξ^i 和余向量 η_j. 可作其乘积, 即张量 $T^i_j = \xi^i\eta_j$, 其为 (1,1) 型, 而迹为 $T^i_i = \xi^i\eta_i$. 这是个标量, 即向量与余向量的内积.

17.2 如果已给出向量 ξ^i 和线性算子 A^k_l, 则乘积 $T^{ik}_l = A^k_l\xi^i$ 为 (2,1) 型的向量. 缩并

$$\eta^k = T^{ik}_i = A^k_i\xi^i$$

仍是个向量, 即对原来向量 ξ^i 应用算子 A^k_l 的结果.

注 利用所引进的向量和余向量的内积, 我们对每个向量指定一个作用于函数的一阶线性微分算子; 因为 $\left(\dfrac{\partial f}{\partial x^i}\right)$ 为余向量 (函数的梯度), 表达式

$$\partial_\xi f = \xi^i\frac{\partial f}{\partial x^i} \tag{14}$$

为标量 (函数沿方向 ξ 的导数). 特别地, 如果 $e_1, \cdots, e_n$ 为向量空间的基, 其中向量 e_k 的坐标等于 $(e_k)^i = \delta^i_k$, 于是由公式 (14) 得到

$$\partial_{e_k}(f) = \frac{\partial f}{\partial x^k}. \tag{15}$$

所以在我们的对应下, 场的基向量转化为 $\dfrac{\partial}{\partial x^1}, \cdots, \dfrac{\partial}{\partial x^k}$. 这样, 向量 ξ 对应于微分算子

$$\partial_\xi = \xi^i\frac{\partial}{\partial x^i}. \tag{16}$$

习题

17.1. 验证置换上和下的指标 $T^{\cdots i_k\cdots}_{\cdots j_l\cdots} \to T^{\cdots j_l\cdots}_{\cdots i_k\cdots}$ 不是一个张量运算.(举例)

17.2. 如果一个 2 阶张量所对应的矩阵是非退化的, 则称其为非退化张量. 证明非退化二阶张量的逆矩阵也是个张量.

§18. (0,k) 型张量

1. 表下指标张量为微分形式

先考虑 (0,1) 型的张量, 即余向量. 我们在前面有过余向量的例子: 函数 f 的梯度 $\left(\dfrac{\partial f}{\partial x^i}\right)$. 回忆在分析中, 我们称表达式

$$df = \frac{\partial f}{\partial x^i}dx^i \tag{1}$$

为函数的微分. 如果给出变换 $x^i = x^i(x^{1'}, \cdots, x^{n'})$, 则

$$dx^i = \frac{\partial x^i}{\partial x^{i'}} dx^{i'}, \tag{2}$$

$$df = \frac{\partial f}{\partial x^i} dx^i = \left(\frac{\partial f}{\partial x^i}\frac{\partial x^i}{\partial x^{i'}}\right) dx^{i'} = \frac{\partial f}{\partial x^{i'}} dx^{i'}, \tag{3}$$

于是表达式 df 对于坐标变换不变. 类比地, 如果我们对任意余向量 (T_i) 给出对应的表达式 $T_i dx^i$(微分形式), 则此表达式对于坐标变换不变.

这个符号 dx^i 是什么? 场的基余向量 e^i 按规则

$$e^i = \frac{\partial x^i}{\partial x^{i'}} e^{i'}, \quad T_i e^i = T_{i'} e^{i'} \tag{4}$$

变换. 最后面的那个等式表明 T_i(相应地 $T_{i'}$) 为在基 $e^1, \cdots, e^n$(相应地 $e^{1'}, \cdots, e^{n'}$) 下余向量的坐标.

我们看到基余向量 e^i 按与 dx^i 一样的规则变换

$$e^i = \frac{\partial x^i}{\partial x^{i'}} e^{i'} \leftrightarrow dx^i = \frac{\partial x^i}{\partial x^{i'}} dx^{i'},$$

$$e^i \leftrightarrow dx^i, \quad e^{i'} \leftrightarrow dx^{i'}. \tag{5}$$

可以说, 符号 dx^i 就是基余向量 e^i. 微分形式 $T_i dx^i$ 对应于余向量按基的展开式 $T_i e^i$. 在上一节我们已看到, 任意余向量是在向量上的线性形式. 特别, 线性形式 $df = \dfrac{\partial f}{\partial x^i} dx^i$ 在向量 $\Delta\xi = \Delta x^i e_i$ 上的取值按定义等于

$$\left(\frac{\partial f}{\partial x^i} dx^i, \Delta\xi\right) = \frac{\partial f}{\partial x^i} \Delta x^i. \tag{6}$$

我们已经知道这个表达式被称为函数 f 的增量的线性主部, 这个增量是由位移向量 $\Delta\xi$ 产生的.

考虑第二个重要情形:(0,2) 型的张量. 这类张量的空间的基由乘积

$$e^i \otimes e^j \tag{7}$$

组成. 按基 (7) 展开任意张量 (T_{ij}) 有形式

$$T_{ij} e^i \otimes e^j. \tag{8}$$

张量 (T_{ij}) 是向量上的双线性形式. 事实上,如果 ξ, η 为两个向量, 则表达式

$$T_{ij} \xi^i \eta^j \tag{9}$$

为标量, 是双线性形式 T 在向量 ξ, η 上的值.

任意一个 (0,2) 型张量分解为对称和反对称 ($T_{ji}=-T_{ij}$) 的和. 这由下面恒等式可清楚看出

$$T_{ij}=T_{ij}^{\mathrm{sym}}+T_{ij}^{\mathrm{alt}},$$

$$T_{ij}^{\mathrm{sym}}=\frac{1}{2}(T_{ij}+T_{ji});\quad T_{ij}^{\mathrm{alt}}=\frac{1}{2}(T_{ij}-T_{ji}). \tag{10}$$

由 (10) 得到对称张量空间的基为

$$\frac{e^i\otimes e^j+e^j\otimes e^i}{2},\quad i\leqslant j, \tag{11}$$

和反称张量空间的基

$$e^i\wedge e^j=e^i\otimes e^j-e^j\otimes e^i, i<j. \tag{12}$$

如果张量 T_{ij} 对称, 则其按基 (11) 的展式为

$$\begin{aligned}T_{ij}e^i\otimes e^j&=\sum_{i\leqslant j}T_{ij}e^i\otimes e^j+\sum_{i>j}T_{ij}e^i\otimes e^j\\&=\sum_{i}T_{ii}e^i\otimes e^i+\sum_{i<j}2T_{ij}\left(\frac{e^i\otimes e^j+e^j\otimes e^i}{2}\right).\end{aligned} \tag{13}$$

如果张量 T_{ij} 为反称, 则其按基 (12) 的展式为

$$T_{ij}e^i\otimes e^j=\sum_{i<j}T_{ij}e^i\otimes e^j+\sum_{i>j}T_{ij}e^i\otimes e^j=\sum_{i<j}T_{ij}e^i\wedge e^j. \tag{14}$$

基 (11) 在微分形式下可记为 $dx^idx^j=dx^jdx^i$, 而基 (12) 可记为 $dx^i\wedge dx^j=-dx^j\wedge dx^i$.

例 18.1. 黎曼度量 g_{ij} 是对称的 (0,2) 型张量的例子. 它按基 dx^idx^j 的展式为

$$g_{ij}dx^idx^j. \tag{15}$$

这是长度元平方的已知公式.

2. (0,k) 型反称张量

定义 18.1. 称张量 $(T_{i_1\cdots i_k})$ 为 $(0,k)$ 型反称张量是说它满足

$$T_{\sigma(i_1,\cdots,i_k)}=\mathrm{sgn}(\sigma)T_{i_1\cdots i_k}. \tag{16}$$

其中 $\mathrm{sgn}(\sigma)=\pm1$ 是置换 σ 的符号; 我们在 §17 的公式 (8) 中定义了它在数组 $(i_1,\cdots,i_k)$ 上的作用.

由于置换指标运算的张量性质, 这个定义不依赖于坐标系的选取. 因此, 张量 $(T_{i_1\cdots i_k})$ 在任意指标的奇置换下改变了符号, 而在指标的偶置换下数值不变.

注　如果阶 k 大于空间的维数 n, 那么反称张量 $(T_{i_1\cdots i_k})$ 恒为零 (显然会有一对重复的指标).

对于反称张量而言使用微分形式的语言非常方便. 这种张量空间的基由

$$dx^{i_1}\wedge\cdots\wedge dx^{i_k},\quad i_1<\cdots<i_k \tag{17}$$

组成, 其中

$$dx^{i_1}\wedge\cdots\wedge dx^{i_k}=\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)e^{\sigma(i_1)}\otimes\cdots\otimes e^{\sigma(i_k)} \tag{18}$$

(对全部指标的置换取和). 如同前一小节中那样, 我们对于反称张量 $(T_{i_1\cdots i_k})$ 得到对应的微分形式

$$T_{i_1\cdots i_k}e^{i_1}\otimes\cdots\otimes e^{i_k}=\sum_{i_1<\cdots<i_k}T_{i_1\cdots i_k}dx^{i_1}\wedge\cdots\wedge dx^{i_k}. \tag{19}$$

表达式 $dx^{i_1}\wedge\cdots\wedge dx^{i_k}$ 对于指标置换为反称的

$$dx^{\sigma(i_1)}\wedge\cdots\wedge dx^{\sigma(i_k)}=\operatorname{sgn}(\sigma)dx^{i_1}\wedge\cdots\wedge dx^{i_k}. \tag{20}$$

例 18.2.　考虑在 n 维空间中的 $(0,n)$ 型反称张量. 每个这样的张量 $(T_{i_1\cdots i_n})$ 由一个数 $T_{12\cdots n}$ 确定. 事实上,

$$T_{\sigma(1,\cdots,n)}=\operatorname{sgn}(\sigma)T_{12\ldots n}, \tag{21}$$

而如果在指标 $i_1,\cdots,i_k$ 中即使有两个相同, 则分量 $T_{i_1\ldots i_k}$ 为零. 因此, 我们有唯一的基的张量, 即 $dx^1\wedge\cdots\wedge dx^n$.

在物理的文献中这个张量的分量以 $\varepsilon_{i_1\cdots i_n}$ 表示. 显见, 分量 $\varepsilon_{i_1\cdots i_k}$ 不为零仅当在指标 $i_1,\cdots,i_n$ 中没有重复的. 于是

$$\varepsilon_{i_1\cdots i_n}=\begin{cases}+1 & 当\ \operatorname{sgn}(i_1,\cdots,i_n)=+1,\\ -1 & 当\ \operatorname{sgn}(i_1,\cdots,i_n)=-1.\end{cases} \tag{22}$$

自然,$T_{i_1\cdots i_n}=T_{12\cdots n}\varepsilon_{i_1\cdots i_n}$.

在坐标变换下 n 阶反称张量如何变化?

定理 18.1.　阶等于空间维数的反称张量在坐标变换 $x^{i'}=x^{i'}(x^1,\cdots,x^n)$ 下其变换规则是

$$T_{12\cdots n}=T_{1'2'\ldots n'}J, \tag{23}$$

其中 $J=\det\left(\dfrac{\partial x^{i'}}{\partial x^j}\right)$ 为变换的雅可比.

证明 我们对基张量 $\varepsilon_{i_1\cdots i_n}$ 来验证公式 (23). 我们有

$$\varepsilon_{12\cdots n}=\varepsilon_{i'_1\cdots i'_n}\frac{\partial x^{i'_1}}{\partial x^{i_1}}\cdots\frac{\partial x^{i'_n}}{\partial x^{i_n}}=\sum_{\sigma\in S_n}\mathrm{sgn}(\sigma)\frac{\partial x^{i'_1}}{\partial x^1}\cdots\frac{\partial x^{i'_n}}{\partial x^n},$$

其中 σ 为置换 $\begin{pmatrix}1 & \cdots & n\\ i'_1 & \cdots & i'_n\end{pmatrix}$. 这个公式是行列式 J 的定义. □

例 18.3. 设 (0,2) 型张量 (g_{ij}) 是非退化二次型, 且 $g=\det(g_{ij})$. 在坐标变换 $x^{i'}=x^{i'}(x^1,\cdots,x^n)$ 下分量 g_{ij} 的变换规则为

$$g_{i'j'}=g_{ij}\frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^j}{\partial x^{j'}},$$

或者写为矩阵形式为 $G'=A^{\mathrm{T}}GA, A=\left(\dfrac{\partial x^i}{\partial x^{i'}}\right), G=(g_{ij}), G'=(g_{i'j'})$. 那么行列式 $g=\det\ G$ 变换规则为

$$g'=\det\ G'=\det(A^{\mathrm{T}}GA)=(\det\ A)^2\det\ G. \tag{24}$$

因此,$\sqrt{|g'|}=\sqrt{|g|}|\det\ A|$.

这样, 我们证明了

推论 表达式 $\sqrt{|g|}dx^1\wedge\cdots\wedge dx^n$ 在 $J=\det\ A=\det\left(\dfrac{\partial x^i}{\partial x^{j'}}\right)>0$ 时是在这种坐标变换下的张量.

复的情形. 如果给出了复空间, 其坐标为 $z^1,\cdots,z^n;\bar z^1,\cdots,\bar z^n$, 其中 $z^\alpha=x^\alpha+iy^\alpha, \bar z^\alpha=x^\alpha-iy^\alpha$, 那么微分形式可以写成

$$T=\sum T^{(p,q)},\quad p+q=k,$$

其中的项

$$T^{(p,q)}=\sum_{\substack{i_1<\cdots<i_p\\ j_1<\cdots<j_q}}T_{i_1\cdots i_p,j_1\cdots j_q}dz^{i_1}\wedge\cdots\wedge dz^{i_p}\wedge d\bar z^{j_1}\wedge\cdots\wedge d\bar z^{j_q} \tag{25}$$

被称做 "(p,q)型的形式".

例如, 设给出形式

$$\Omega=T_{\alpha\beta}dz^\alpha\wedge d\bar z^\beta,$$

其中 $T_{\beta\alpha}=-\overline{T}_{\alpha\beta}$. 于是矩阵 $(iT_{\alpha\beta})$ 满足条件

$$iT_{\beta\alpha}=\overline{iT}_{\alpha\beta}$$

从而是埃尔米特形式 $\sum iT_{\alpha\beta}dz^\alpha d\bar z^\beta$ 的矩阵. 于是, 在复情形下埃尔米特度量为 (1,1) 型的形式.

3. 微分形式的外积. 外代数

作为在 §17 中引进的张量的代数运算的应用, 我们来定义两个反称张量的外积, 它们分别是 $(0,p)$ 型和 $(0,q)$ 型 (相应地为 p 和 q 阶的微分形式). 假设

$$\begin{aligned} \omega_1 &= \sum_{i_1<\cdots<i_p} T_{i_1\cdots i_p} dx^{i_1}\wedge\cdots\wedge dx^{i_p}, \\ \omega_2 &= \sum_{j_1<\cdots<j_q} S_{j_1\cdots j_q} dx^{j_1}\wedge\cdots\wedge dx^{j_q}. \end{aligned} \tag{26}$$

定义 $p+q$ 阶形式 ω 为

$$\omega=\omega_1\wedge\omega_2=\sum_{k_1<\cdots<k_{p+q}} R_{k_1\cdots k_{p+q}} dx^{k_1}\wedge\cdots\wedge dx^{k_{p+q}}, \tag{27}$$

其中令

$$R_{k_1\cdots k_{p+q}}=\sum_{\sigma\in S_{p+q}}\frac{\mathrm{sgn}(\sigma)}{p!q!}T_{\sigma(k_1)\cdots\sigma(k_p)}S_{\sigma(k_{p+1})\cdots\sigma(k_{p+q})}. \tag{28}$$

数 $R_{k_1\cdots k_{p+q}}$ 构成了张量, 它是由张量 $(T_{i_1\cdots i_p}),(S_{j_1\cdots j_q})$ 的张量积和指标置换的组合得到. 由构造知此张量为反称, 从而定义 (28) 不依赖于坐标的选取.

引理 18.1. 如果 ω_1 为 p 阶形式,ω_2 为 q 阶形式, 则微分形式的外积是个双线性可结合的算子, 并且

$$\omega_2\wedge\omega_1=(-1)^{pq}\omega_1\wedge\omega_2. \tag{29}$$

证明 双线性是显见的. 结合性则直接由公式 (28) 验证. 公式 (29) 由置换

$$\begin{pmatrix} 1 & \cdots & q & q+1 & \cdots & p+q \\ p+1 & \cdots & p+q & 1 & \cdots & p \end{pmatrix}$$

的符号为 $(-1)^{pq}$(验证之!) 得到. 引理得证. □

注 不难验证, 对基形式 $\omega_1=dx^i,\omega_2=dx^j$ 的外积 (28) 与在第 2 小节中定义的记号 $dx^i\wedge dx^j$ 相同. 一般地, 如果 $\omega_1=dx^{i_1}\wedge\cdots\wedge dx^{i_p},\omega_2=dx^{j_1}\wedge\cdots\wedge dx^{j_q}$, 则

$$\omega_1\wedge\omega_2=dx^{i_1}\wedge\cdots\wedge dx^{i_p}\wedge dx^{j_1}\wedge\cdots\wedge dx^{j_q}. \tag{30}$$

外积给出了在所给点处的形式的空间中一个 "外代数" 结构

4. $(k,0)$ 型反称张量 (多向量). 对于反交换变量的积分

常常称具上指标的反称张量为多向量. 用外代数的语言描述它们也是方便的. 如果 $e_1,\cdots,e_n$ 为标准的基向量, 则类似于 (17),(18) 给出了 $(k,0)$ 型反称张量 (多向量) 的基为 $e_{i_1}\wedge\cdots\wedge e_{i_k},i_1<\cdots<i_k$, 在这里的与基形式 $(e^j\leftrightarrow dx^j)$ 的缩并为

$$(e_{i_1}\wedge\cdots\wedge e_{i_k},e^{j_1}\wedge\cdots\wedge e^{j_k})=\delta^{j_1}_{i_1}\cdots\delta^{j_k}_{i_k}, \tag{31}$$

$j_1<\cdots<j_k$. 与 $(0,k)$ 型张量完全类似, 多向量构成了外代数, 它由符号 $e_1,\cdots,e_n$ 生成. 成立下面的类似于定理 18.1 的定理.

定理 18.2. 在坐标变换下成立公式

$$T^{12\cdots n}=J^{-1}T^{1'2'\cdots n'},\tag{32}$$

其中 $J=\det\left(\dfrac{\partial x^{i'}}{\partial x^j}\right)$ 为变换的雅可比.

证明 与定理 18.1 的证明没有区别. □

依照现代的关于量子场论的文献, 我们介绍 "关于反交换变量 $\theta_1,\cdots,\theta_n$ 的积分":

$$\int\underset{\mathbb{R}^n}{\cdots}\int f(\theta_1,\cdots,\theta_n)d\theta_1\wedge\cdots\wedge d\theta_n;\tag{33}$$

这里的 $f(\theta_1,\cdots,\theta_n)$ 为在由 $\theta_1,\cdots,\theta_n$ 生成的外代数中的多项式, 而积分须取在全空间 (不是在区域上!). 在 θ 空间中的线性变换下, 量 $d\theta_1\wedge\cdots\wedge d\theta_n$ 按定义其变换为

$$d\theta_1\wedge\cdots\wedge d\theta_n=J^{-1}d\theta'_1\wedge\cdots\wedge d\theta'_n,\tag{34}$$

其中 $J=\det\left(\dfrac{\partial\theta_i}{\partial\theta'_j}\right)$ 为普通的数. 积分规则定义为

$$\int_{\mathbb{R}}\theta_j d\theta_j=1,\quad \int_{\mathbb{R}}d\theta_j=0,\tag{35}$$

$$\iint_{\mathbb{R}^2}f(\theta_1)\wedge g(\theta_2)d\theta_1\wedge d\theta_2=\left(\int_{\mathbb{R}}f(\theta_1)d\theta_1\right)\wedge\left(\int_{\mathbb{R}}g(\theta_2)d\theta_2\right)\tag{36}$$

(其他的规则是其自然推广), 其中 $\theta_1,\theta_2,\cdots,\theta_n$ 为任意一组外代数生成元. 由公式 (31),(32) 得到结论: 在做对照 $d\theta_j\to e_j,\theta_j\to e^j$ 后,$d\theta_1\wedge\cdots\wedge d\theta_n$ 成为 $(n,0)$ 型张量即多向量的基, 而 $f(\theta_1,\cdots,\theta_n)$ 的首项成为 $(0,n)$ 型张量, 而 "积分" 则是它们通常的缩并. 这样一来, 对反交换变量的积分就是通常的张量缩并, 而这些张量在嘉当外代数的公式化描述中既作为形式也作为多向量. 作这种解释的用处表现于下面的事实.

习题 18.1. 设 $f(\theta_1,\cdots,\theta_n)=\exp(\frac{1}{2}a^{ij}\theta_i\wedge\theta_j),a^{ji}=-a^{ij}$.
证明公式

$$\int\underset{\mathbb{R}^n}{\cdots}\int f(\theta_1,\cdots,\theta_n)d\theta_1\wedge\cdots\wedge d\theta_n=\sqrt{\det(a^{ij})}.\tag{37}$$

在空间 $\mathbb{R}^n$ 上的表达式 (33) 给出了 (n,n) 型张量, 它分别对上和下指标为反称. 这样的张量必为 $\mathbb{R}^n$ 上的标量.

在积分 (33) 中变换 $\theta(\theta')$ 可以是非线性的. 唯一的要求是它要保持所谓的 $\mathbb{Z}_2$ 分次不变, 就是说多项式 $\theta(\theta')$ 应该是奇数阶表达式在外代数中的线性组合.

例 18.4.　设 $\theta=(\theta_1,\theta_2,\theta_3), f(\theta)=\theta_1$. 于是由定义有

$$\int f(\theta)d\theta_1\wedge d\theta_2\wedge d\theta_3=0.$$

考虑变换

$$\theta_1=\theta_1'+\theta_1'\wedge\theta_2'\wedge\theta_3',\quad \theta_2=\theta_2',\quad \theta_3=\theta_3'.$$

以及此变换的雅可比矩阵 (准确的定义可见下面):

$$\left(\frac{\partial\theta_i}{\partial\theta_j'}\right)=\begin{pmatrix}1+\theta_2'\wedge\theta_3' & -\theta_1'\wedge\theta_3' & \theta_1'\wedge\theta_2'\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

雅可比 $J=\det\left(\frac{\partial\theta_i}{\partial\theta_j'}\right)$ 具有形式 $J=1+\theta_2'\wedge\theta_3'; J^{-1}=1-\theta_2'\wedge\theta_3'$. 验证下面的等式:

$$\int f(\theta(\theta'))J^{-1}d\theta_1'\wedge d\theta_2'\wedge d\theta_3'=0.$$

更一般地, 如果

$$\theta_i=a_{ij}(\theta_1',\cdots,\widehat{\theta_j'},\cdots,\theta_n')+\theta_j'\wedge b_{ij}(\theta_1',\cdots,\widehat{\theta_j'},\cdots,\theta_n'),$$

则我们有

$$\frac{\partial\theta_i}{\partial\theta_j'}=b_{ij}(\theta_1',\cdots,\widehat{\theta_j'},\cdots,\theta_n').$$

显然, 在 $\mathbb{Z}_2$ 分次变换下雅可比矩阵的元素可交换.

习题 18.2.　证明下面的变量变换公式:

$$\int f(\theta(\theta'))J^{-1}d\theta_1'\wedge\cdots\wedge d\theta_n'=\int f(\theta)d\theta_1\wedge\cdots\wedge d\theta_n, \tag{38}$$

其中 $J=\det\left(\frac{\partial\theta_i}{\partial\theta_j'}\right)$.

提示　如果 $f(\theta)$ 是最高阶的单项式, 则 (38) 显然成立. 需要分别分析在多项式 $f(\theta)$ 中较低阶的项, 它们在变换后只贡献了零值, 这就如同上面例子中所证明的那样.

习题

18.3.　设 $\omega^j=a_i^j dx^i$. 证明公式

$$\omega^{i_1}\wedge\cdots\wedge\omega^{i_k}=J^{i_1\cdots i_k}_{j_1\cdots j_k}dx^{j_1}\wedge\cdots\wedge dx^{j_k},$$

其中 $J^{i_1\cdots i_k}_{j_1\cdots j_k}$ 为矩阵 (a^j_i) 的子式, 它由 $i_1,\cdots,i_k$ 行和 $j_1,\cdots,j_k$ 列相交部分构成. 特别,

$$\omega^1\wedge\cdots\wedge\omega^n=\det(a^j_i)dx^1\wedge\cdots\wedge dx^n.$$

18.4. 求 (在给定点上)k 形式空间的维数.

18.5. 证明 $\sum\limits_k a^j_{i_k}\varepsilon_{i_1\cdots i_{k-1}ji_{k+1}\cdots i_n}=\varepsilon_{i_1\cdots i_n}\mathrm{Tr}(a^j_i)$, 其中 $\mathrm{Tr}(a^j_i)=a^i_i$ 为矩阵的迹.

§19. 黎曼和伪黎曼空间中的张量

1. 升标和降标

设 g_{ij} 为 (0,2) 型张量, 它给出了黎曼或伪黎曼度量. 我们已经知道, 如果给出了两个向量 ξ^i,η^j, 则可定义它们的内积:

$$\langle\xi,\eta\rangle=\xi^i\eta^jg_{ij}. \tag{1}$$

(我们在这里用到了张量的乘积和缩并运算.)

类比地, 如果 g^{ij} 为 (2,0) 型张量, 则它给出了余向量 ξ_j,η_i 的内积, 其公式为

$$\langle\xi,\eta\rangle=g^{ij}\xi_j\eta_i. \tag{2}$$

有了度量 g_{ij} 则可以定义非常重要的降标运算. 如果 $T^{i_1\cdots i_p}_{j_1\cdots j_q}$ 为 (p,q) 型张量, 则令

$$T^{i_2\cdots i_p}_{i_1j_1\cdots j_q}=g_{i_1k}T^{ki_2\cdots i_p}_{j_1\cdots j_q} \tag{3}$$

可以构造出一个 $(p-1,q+1)$ 型的向量 $T^{i_2\cdots i_p}_{i_1j_1\cdots j_q}$.

容易看出这仍是一个张量 (是乘以张量 g_{ij} 的运算和缩并的迭代).

称由张量 $T^{i_1\cdots i_p}_{j_1\cdots j_q}$ 到张量 $T^{i_2\cdots i_p}_{i_1j_1\cdots j_q}$ 的过程为用度量 g_{ij} 降标i_1.

例 19.1. 如果 ξ^i 为向量, 在降标后我们则得到了余向量

$$\xi_i=g_{ij}\xi^j. \tag{4}$$

因此降标给出了向量空间到余向量空间的映射. 这个对应关系可以以下面的方式描述: 如果 $\langle,\rangle$ 为对应于 g_{ij} 的内积, 则向量 η 对应于线性形式 (余向量), 它在向量 ξ 的取值为 $\langle\xi,\eta\rangle$.

反之, 在原来的度量 g_{ij} 下为了提升下指标则必须考虑逆度量, 即那样的 (g^{ij}) 使得

$$g^{ij}g_{jk}=\delta^i_k \tag{5}$$

(记住有 $\det(g_{ij})\neq0$).

作为定义我们有升标

$$T^{j_1 i_1 \cdots i_p}_{j_2 \cdots j_q} = g^{j_1 k} T^{i_1 \cdots i_p}_{k j_2 \cdots j_q}. \tag{6}$$

引理 19.1. 如果我们降标, 然后再升标, 则我们得到原来的张量.

证明 对张量 $T^{i_1 \cdots i_p}_{j_1 \cdots j_q}$ 降指标 i_1, 我们得到张量 $g_{i_1 k} T^{k i_2 \cdots i_p}_{j_1 \cdots j_q}$. 之后再升指标 i_1, 我们得到张量

$$g^{i_1 l} g_{lk} T^{k i_2 \cdots i_p}_{j_1 \cdots j_q} = \delta^{i_1}_k T^{k i_2 \cdots i_p}_{j_1 \cdots j_q} = T^{i_1 \cdots i_p}_{j_1 \cdots j_q},$$

与原来的重合. 证完. □

在前面我们已经看到, 矩阵 (g^{ij}) 给出了余向量的内积. 因此, 在给出了度量 g_{ij} 后我们便定义了余向量的内积 g^{ij}. 这个余向量的内积在下面的条件下是唯一确定的, 即在升标运算后我们得到与向量的内积相同的内积. 这可由下面的论断得到.

引理 19.2. 成立下面等式: 对两个向量 $\xi = (\xi^i), \eta = (\eta^i)$ 和对应的两个余向量 $\widehat{\xi} = (\widehat{\xi}_i) = (g_{ij}\xi^j), \widehat{\eta} = (\widehat{\eta}_i) = (g_{ij}\eta^j)$, 它们的内积相同: $\langle \widehat{\xi}, \widehat{\eta} \rangle = \langle \xi, \eta \rangle$.

证明 因为 $\langle \xi, \eta \rangle = g_{ij}\xi^i\eta^j, \langle \widehat{\xi}, \widehat{\eta} \rangle = g^{ij}\widehat{\xi}_j\widehat{\eta}_i$, 则

$$\langle \widehat{\xi}, \widehat{\eta} \rangle = g^{ij}\widehat{\xi}_j\widehat{\eta}_i = g^{ij} g_{jk}\xi^k g_{il}\eta^l = \delta^i_k \xi^k g_{il}\eta^l = \xi^i g_{ik}\eta^k = \langle \xi, \eta \rangle.$$

证完. □

注 不难看出, 我们目前只用到非退化的张量 g_{ij}, 即 $g = \det(g_{ij}) \neq 0$; 正定性 (还有对称性) 在上述论断及公式中并没起到作用. 在第 5 章中我们将对与哈密顿形式体系相关的反称度量作类似的构造.

2. 二次型的特征值

假设给出了一个线性算子 T^j_i, 即 (1,1) 型的张量. 如果没有度量, 那么谈及此张量的对称性或者反称性就没有意义, 这是因为不能在其指标 i 和 j 之间进行置换.

如果有了度量 g_{ij}, 则可降标:$T_{ij} = g_{ik}T^k_j$; 这时我们得到了 (0,2) 型张量, 它可定义在两个向量上的双线性形式 $\{,\}_T$: 如果 $\langle , \rangle$ 为内积, 它对应于 g_{ij}, 则这个双线性形式为

$$\{\xi, \eta\}_T = \langle \xi, T\eta \rangle = \xi^i g_{ij} T^j_k \eta^k, \tag{7}$$

$$\xi = (\xi^i), \quad \eta = (\eta^i).$$

定义 19.1. 称在具度量 g_{ij} 的空间中的线性算子 T^j_i 为对称的 (反称的) 是说双线性形式 $T_{ij} = g_{ik}T^k_j$ 对称,$T_{ji} = T_{ij}$(反称,$T_{ji} = -T_{ij}$).

定理 19.1. 在具黎曼度量或伪黎曼度量 g_{ij} 的空间中线性算子 $T = (T^i_j)$ 为对称或反称的充要条件是对任意向量 $\xi = (\xi^i)$ 和 $\eta = (\eta^i)$ 满足关系

$$\langle T\xi, \eta \rangle = \langle \xi, T\eta \rangle \ (对称), \tag{8}$$

$$\langle T\xi, \eta \rangle = -\langle \xi, T\eta \rangle \ (反称). \tag{9}$$

由公式 (7) 得知, 定理显然成立.

现设在黎曼或伪黎曼空间中给出了对称的双线性形式 T_{ij}. 于是经升标可考察算子 $T^i_j = g^{ik}T_{kj}$.

定义 19.2. 称算子 $T^i_j = g^{ik}T_{kj}$ 的特征值为二次型 T_{ij} 在度量 g_{ij} 下的特征值.

如果 λ 为算子 T^i_j 的特征值, ξ^i 为对应的特征向量, 那么, 由定义有

$$T^i_j\xi^j = \lambda\xi^i \Longleftrightarrow g^{ik}T_{kj}\xi^j = \lambda\xi^i. \tag{10}$$

因此特征向量 ξ^i 为方程组

$$T_{kj}\xi^j = \lambda g_{ki}\xi^i, \quad k = 1, \cdots, n \tag{11}$$

的解 (比较方程组 (10)).

线性算子 $T^i_j = g^{ik}T_{kj}$ 的迹 $\mathrm{Tr}\, T^i_j$ 和行列式 $\det T^i_j$ 是形式 T_{ij} 的度量的不变量 (即依赖于度量的在坐标变换下不变的量). 特别, 迹有形状

$$T^i_i = g^{ik}T_{ik}. \tag{12}$$

例 19.2. 空间 $\mathbb{R}^3$ 的曲面 $r = r(u, v), r = (x, y, z)$ 上取两个二次型 $(x^1 = u, x^2 = v)$:

1) 度量 $g_{ij}dx^idx^j$ 为张量 g_{ij};

2) 第二基本型 $b_{ij}dx^idx^j$ 为张量 b_{ij}(因为曲面是 2 维的, 这里是对从 1 到 2 的指标双重取和).

在 §8 中得到过公式

$$\text{高斯曲率} \quad K = \frac{\det(b_{ij})}{\det(g_{ij})},$$

$$\text{平均曲率} \quad H = b^i_i = g^{ij}b_{ij}.$$

因此, 平均曲率是在度量 g_{ij} 存在情形下二维张量 b_{ij} 的迹. 对于高斯曲率而言, 我们有

$$K = (\det(g_{ij}))^{-1}\det(b_{ij}) = \det(g^{ik}b_{kj}) = \det(b^i_j).$$

从而高斯曲率也是第二基本型的度量的不变量.

3. * 算子

在存在度量 g_{ij} 的情形下, 可以给出运算 *, 使得 $(0, k)$ 型反称张量与 $(0, n-k)$ 型反称张量等同.

定义 19.3. 如果 T 为 $(0, k)$ 型反称张量, 其分量为 $T_{i_1\cdots i_k}$, 则以 $*T$ 表示一个 $(0, n-k)$ 型反称张量

$$(*T)_{i_{k+1}\cdots i_n} = \frac{1}{k!}\sqrt{|g|}\varepsilon_{i_1\cdots i_n}T^{i_1\cdots i_k}, \tag{13}$$

其中 $T^{i_1\cdots i_k}$ 为对应于张量 $T_{i_1\cdots i_k}$ 的一个 $(k,0)$ 型张量

$$T^{i_1\cdots i_k}=g^{i_1j_1}\cdots g^{i_kj_k}T_{j_1\cdots j_k}. \tag{14}$$

在 §18.2 中我们知道 $\sqrt{|g|}\varepsilon_{i_1\cdots i_n}$ 是 n 维空间上的一个对于具正雅可比的坐标变换的张量. 因此 $*T$ 也是对于这种变换的张量, 而它的反称性则是显见的.

注 成立公式 (请验证)

$$*(*T)=(-1)^{k(n-k)}\mathrm{sgn}(g)T. \tag{15}$$

4. 欧氏空间的张量

在欧氏空间中, 欧氏坐标下的度量 g_{ij} 具有形状 $g_{ij}=\delta_{ij}$. 因此在降 (升) 标时任意一个张量的分量都没有改变:

$$T^{i_2\cdots i_p}_{i_1j_1\cdots j_q}=\delta_{i_1i}T^{ii_2\cdots i_p}_{j_1\cdots j_q}=T^{i_1i_2\cdots i_p}_{j_1\cdots j_q}. \tag{16}$$

那么, 在欧氏坐标下, 上指标和下指标没有区别; 譬如, 可以认为所有的指标都是下指标. 这种记号在保持欧氏度量不变的变换下为不变的; 而这类变换就是正交变换和平移.

特别, 在欧氏坐标下, 算子的矩阵和对应于它的二次型的矩阵相同. 梯度的分量的变换在欧氏空间的运动下就像向量变换一样, 等等.

我们考虑在具欧氏坐标 x,y,z 的三维欧氏空间中算子 $*$ 的作用:

$$*dx=dy\wedge dz,\quad *dy=-dx\wedge dz,\quad *dz=dx\wedge dy$$

(请验证). 任意一个 1 形式 (余向量) 为 $\omega=Pdx+Qdy+Rdz$. 对此形式有

$$*\omega=Pdy\wedge dz+Qdz\wedge dx+Rdx\wedge dy,$$
$$*(*\omega)=\omega.$$

如果 f 为标量, 则 $*f=fdx\wedge dy\wedge dz$ 是个 3 阶形式, 而且 $*(fdx\wedge dy\wedge dz)=f$.

习题

19.1. 设 $\varepsilon_{\alpha\beta\gamma}$ 是在前面 (见 §18.2) 已考虑过的反称张量, 但是在三维欧氏空间中. 证明下面的公式:

a) $\varepsilon_{\alpha\beta\gamma}\varepsilon_{\lambda\mu\nu}=\begin{vmatrix}\delta_{\alpha\lambda} & \delta_{\alpha\mu} & \delta_{\alpha\nu}\\ \delta_{\beta\lambda} & \delta_{\beta\mu} & \delta_{\beta\nu}\\ \delta_{\gamma\lambda} & \delta_{\gamma\mu} & \delta_{\gamma\nu}\end{vmatrix}$,

b) $\varepsilon_{\alpha\beta\gamma}\varepsilon_{\lambda\mu\nu}=\delta_{\alpha\lambda}\delta_{\beta\mu}-\delta_{\alpha\mu}\delta_{\beta\lambda}$,

c) $\varepsilon_{\alpha\beta\gamma}\varepsilon_{\lambda\beta\gamma} = 2\delta_{\alpha\lambda}$,

d) $\varepsilon_{\alpha\beta\gamma}\varepsilon_{\alpha\beta\gamma} = 6$

(每个等式中均表明对于重复的指标求和);$\delta_{\alpha\beta}$ 为克罗内克符号.

19.2. 设

$$\omega_1 = \sum_{i_1 < \cdots < i_p} T_{i_1 \cdots i_p} dx^{i_1} \wedge \cdots \wedge dx^{i_p},$$

$$\omega_2 = \sum_{j_1 < \cdots < j_p} S_{j_1 \cdots j_p} dx^{j_1} \wedge \cdots \wedge dx^{j_p}.$$

令 $\{\omega_1, \omega_2\} = \frac{1}{p!} g^{i_1 j_1} \cdots g^{i_p j_p} T_{i_1 \cdots i_p} S_{j_1 \cdots j_p}$. 证明

$$\omega_1 \wedge \omega_2 = \{\omega_1, \omega_2\} \sqrt{|g|} dx^1 \wedge \cdots \wedge dx^n.$$

§20. 晶体群和平面与空间旋转群的有限子群. 不变张量的例子

在本节中我们要来了解晶体群. 然后将对正交群 $O(2)$ 和 $O(3)$ 中的有限子群分类.

我们将考虑整个充满三维欧氏空间或欧氏平面的晶体结构 (或平面的结构). 标准的晶体模型之一如下: 假定晶体由某类原子组成, 它们被牢牢地固定在空间 (或平面) 中, 并且按严格定义的方式相互有序地排列着.

定义 20.1. 晶体的所有原子的集合称为晶体格.

我们要极大地缩小所考虑的格, 从而只限于那些平移不变的格, 对此我们现在就转向它的定义. 就是这样的格, 它们对应了真正的物理的晶体 (在某种近似下). 我们认为晶体的格总包含了某个这样的子集, 即由向量 $\alpha = n_1\alpha_1 + n_2\alpha_2 + n_3\alpha_3$ (或在平面上的 $\alpha = n_1\alpha_1 + n_2\alpha_2$) 的端点表示的所有点 (原子) 决定. 这里的 n_1, n_2, n_3 为任意整数, 向量 $\alpha_1, \alpha_2, \alpha_3$ 被称做基本平移向量. 有时也称向量 $\alpha_1, \alpha_2, \alpha_3$ 为这个格的*本原向量*,并总假定它们是线性无关的.

重要的要求条件是: 通常假定晶体的格 R 在按向量 $\alpha_1, \alpha_2, \alpha_3$ 的所有基本的平移下变到自己, 就是说, 在这些基本平移的整数系数的线性组合的平移下变到自己; 也就是要求使晶体结构在所有由向量 $\alpha_1, \alpha_2, \alpha_3$ 生成的平移下保持不变. 这是真正的 (“无限”) 晶体的基本性质之一.

记按向量 $\alpha_1, \alpha_2, \alpha_3$ 的平移分别为 τ_1, τ_2, τ_3. 于是我们所考虑的平移中任一个均可写为

$$T = n_1\tau_1 + n_2\tau_2 + n_3\tau_3.$$

定义 20.2. 如果在形如 $T = n_1\tau_1 + n_2\tau_2 + n_3\tau_3$ 的任意的平移下格 R 变到自己, 则称其为*平移不变*. 对平面格的情形其定义完全类似.

因此, 自此后我们将只考虑平移不变的格(在平面或在空间中).

注　通常叙述晶体的数学理论是从空间 $\mathbb{R}^3$ 的运动群 G_3 的离散子群 Γ 开始, 以使 G_3/Γ 为紧. 称其为晶体群 Γ. 可以证明平移子群在 Γ 中具有有限指数. 通常这是一个最为对称的轨道. 读者可以在书 [46] 中找到详细的叙述.

定义 20.3.　称由向量 $\alpha_1, \alpha_2, \alpha_3$ 构成的平行六边形为 (晶体) 格的基本胞腔.

注　设给出了一个格 R. 我们总是假定向量 $\alpha_1, \alpha_2, \alpha_3$(平面情形的 α_1, α_2) 使得它们张成的平行六面体的体积为尽可能地最小.

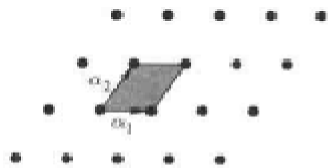


图 18

在图 18 上表示了最简单的二维格; 也指出了它的基本胞腔. 很清楚, 由于晶体的平移不变性, 整个晶体 (格) 便由平移的基本胞腔的并组成. 图 18 表示的最简单平面格具有这样的性质, 使得每个原子 (即格的每个点) 可以由任一个原子经某个平移 $T = n_1\tau_1 + n_2\tau_2$ 得到. 这种情形可以用下面的词语表示: 所有基本平移的集合在格上可迁. 但是这种令人满意的情形远不是对所有的格都成立. 特别, 这种情形可能出于如下原因而发生: 一般说来, 格由某些不同类型的原子组成, 而由于 (在我们的模型的框架内) 自然地要求使得在基本平移下, 一种类型的原子还是转移到同一类的原子, 而不会转移到代表另一类型原子的格子点上去. 因此, 基本平移的集合在格上不是可迁的. 例如在图 19 上展示的这种格.

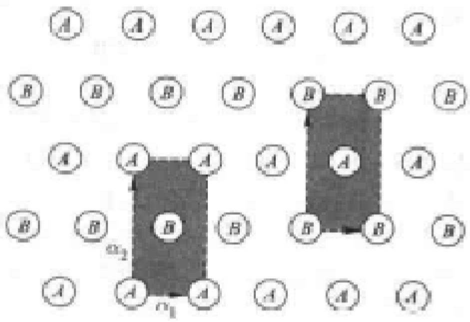


图 19

在这里 A 型的原子不能以基本平移转移到 B 型的原子 (类似地,B 型原子不能转移到 A 型原子). 同时, 基本平移分别在 A 型原子上和在 B 型原子上可迁: 任意 A 型原子 (相应 B 型的) 可以以某个基本平移由任意一个 A 型原子 (相应 B 型的) 得到.

因此我们看到, 为了完全确定地给出晶体格仅给出基本平移的集合是不够的. 另一方面, 我们清楚知道, 由于所有的格是基本胞腔的并, 故而为了完全描述格连同所有基本平移只要给出在任一个基本胞腔中每个原子的位置就够了. 最后, 在所有格的集合中有这样一些格, 它们中基本胞腔的所有原子都可以由这个胞腔的一个原子经基本平移 (在平面中为两个平移, 空间中则为三个) 得到.

定义 20.4.　如果格 R(二维或三维) 的所有原子都位于形如 $n_1\alpha_1 + n_2\alpha_2 + n_3\alpha_3$ (对平面为 $n_1\alpha_1 + n_2\alpha_2$) 的点上, 其中 n_1, n_2, n_3 为任意整数, 则称 R 为布喇菲格.

可以指出, 基本平移的集合在布喇菲格上是可迁的. 由定义显然得到, 不同的布喇菲格相互的差异仅仅在于基本胞腔的形式不同. 特别, 仿射变换可以把两个任意的

布喇菲格互相变换, 就是说, 从仿射的观点看来只有一个布喇菲格. 度量上不同的布喇菲格 (即在正交变换和平行移动下互不重合的格) 以基本平移向量的角度和长度来区分.

定义 20.5. 设 $X_1, \cdots, X_n$ 为所有位于基本胞腔内部的原子; 则说向量 $X_1, \cdots, X_n$(为基本胞腔顶点的原点到点 $X_1, \cdots, X_n$) 构成了格的基 (图 20).

命题 20.1. 给出了基本平移向量和格 R 的基则完全确定了整个 R.

这个命题以显然的方式由我们的基的定义和由格的平移不变性质得到. 把更详细的讨论留给读者, 当作一个关于枚举式定义的初等命题.

我们最终要建立在理想的无限晶体的性质与实际的具有边界的晶体的性质之间对应关系. 在图 21 上展示了理想化的真正三维晶体, 它有边界而无瑕疵. 这里的数 N_1, N_2, N_3 表示在我们的平行六面体 (晶体) 各对应边中能容纳的基本胞腔的个数, 即 $AB = N_1\alpha_1, BC = N_2\alpha_2, AD = N_3\alpha_3$.

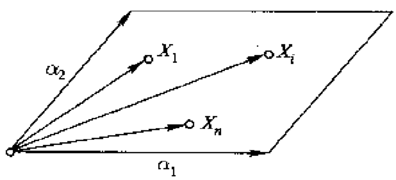


图 20

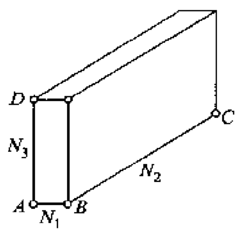


图 21

我们在真正的晶体上, 允许按任意的、可以被向量 $\alpha_1, \alpha_2, \alpha_3$ 整除的向量进行平移; 但我们将以下面的方式进行: 譬如, 格按向量 α_1 位移时, 我们切下晶体向右边缘突出的部分, 并将其粘到左边缘上, 它正好按向量 α_1 深入到平行六面体内部. 事实表明, 使用这样的形式模型是合理的: 在这种 "周期性的" 过程中, 大部分的物理构造和计算都没有变化. 很清楚, 这种观点完全等价于考虑无限的理想晶体. 引进了上面对晶体边缘的周期性条件的描述, 使我们能以直观的几何语言进行解释. 因为晶体的边缘 (譬如, 为简单起见, 我们考虑一维的晶体) 由两个编号为 1 和 N 的原子构成, 则这种 "粘合" 晶体 (即一维的晶体, 满足其端点的周期性条件) 的任意平移化为一个圆按 $2\pi/N$ 倍数角的旋转. 在晶体的平面情形, 周期性条件等价于引进了环面 T^2. 三维晶体的周期性条件粘合成了环 T^3.

与每个格自然关联的概念是对称胞腔 (不要与基本胞腔混淆!). 对称胞腔包含了某个特定的原子作为它的中心.

定义 20.6. 固定格 R 的原子. 如果空间 (在平面格情形时相应地为平面) 的一个集合中的点, 其离我们选取的固定点的距离比到此格的其他原子更近, 则称此集合为 (以给定点为中心的) 对称胞腔. 常称对称胞腔为维涅尔 – 赛茨胞腔(在

离散群论中则称为"狄里克雷区域").

图 22 展示了所谓的六边形二维 (平面) 格, 它标出了基本胞腔和对称胞腔 (这些胞腔不同!). (二维) 对称胞腔的边缘是那些在中点垂直于格的所有连接中心点和最靠近的相邻原子的棱.

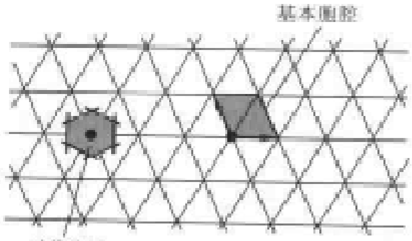


图 22

我们现转向研究保持格不变的变换 (即将其变到自己). 考虑空间 (相应地, 平面) 的运动群, 即所有线性 (非齐次) 变换, 它们保持二次型 $ds^2=(dx^1)^2+(dx^2)^2+(dx^3)^2$ (相应地, $ds^2=(dx^1)^2+(dx^2)^2$) 不变. 我们以 G_3(平面时 G_2) 记此群.

命题 20.2. 群 G_3 (相应地 G_2) 中任意一个元素 g 唯一地表示为两个变换的复合:$g=T\circ\alpha$, 其中 T 为伸缩, 而 α 为旋转 (正常的或不正常的, 即其行列式为 $+1$ 或 -1).

证明　像在 §4.3 中所表明的那样, 三维空间中的任何运动必属于后面两种类型之一:1) 螺旋运动 $g=T\alpha$, α 为旋转, $\det\alpha=+1$, T 为沿旋转轴的平移; 2) $g=\alpha$, 其中 $\alpha\in O(3)$, 且 $\det\alpha=-1$(反射旋转). 在这两种类型中每一种中, 被描述的表示形式唯一地为 $g=T\alpha$ 或 $g=\alpha$. 命题证完. □

注　因为平移和旋转一般并不可换, 故 $T\alpha\neq\alpha T$(请构造例子!)

平移 $\{T\}$ 的集合显然构成 G_3 的子群 (相应地,G_2), 它是 G_3 中 (G_2 中) 的正规子群. 事实上, 变换 gTg^{-1} 对任意旋转 $g\in O(3)$ 仍是平移, 这里的 $O(3)$ 仍像前面那样表示绕点 O 的旋转群 (对平面情形 $g\in O(2)$). 注意, 群 $O(3)$(相应地,$O(2)$) 不是群 G_3 的正规子群 (相应地,G_2).

我们现在在整个变换群 G_3(群 G_2) 中挑出所有使某个固定的格 R 变到自己的变换.

定义 20.7. 群 G_3(群 G_2) 中使得把格 R 变到自己的变换 (运动) 的集合被称做这个格的空间群, 我们以 $G_3(R)$ 表示它 (相应地, $G_2(R)$).

显然, 集合 $G_3(R)$(相应地 $G_2(R)$) 是在通常代数意义下的群.

考虑到在平面上类似的定义和事实, 我们对三维欧氏空间的情形进行进一步的形式化的定义.

群 $G_3(R)$ 包含了子群 $T_3(R)$, 即格的平移群.

定义 20.8. 称由所有可能的平移 T 组成的群 $G_3(R)$ 的子群 $T_3(R)$ 为晶体(即格 R)平移群 (记住, 格 R 的任意平移具形式 $T=n_1\tau_1+n_2\tau_2+n_3\tau_3$, 其中 τ_1,τ_2,τ_3 为格的基本向量 $\alpha_1,\alpha_2,\alpha_3$ 生成的平移).

我们还要记住, 我们总是在考虑平移下不变的格 R.

容易证明, 子群 $T_3(R)$ 是群 $G_3(R)$ 的正规子群. 事实上, 必须证明如果 $g\in$

$G_3(R)$ 及 $t \in T_3(R)$, 则 $gtg^{-1} \in T_3(R)$ 对所有 g 和 t 成立. 换句话说, 需要确信变换 gtg^{-1} 仍然是格的平移. 然而这由格的不变性便得到.

在下文中我们要设定一个属于我们的格中的一个空间的固定点, 即设为 "坐标原点"; 例如, 基本胞腔的顶点为基本平移向量的起点. 我们将考虑对于这个点 O 的格的所有可能的旋转 (行列式为 ± 1).

上面已证任意变换 $g \in G_3$ 有唯一的表示 $g = T\alpha$, 其中 T 为平移,α 为行列式是 ± 1 的旋转. 有时为了有效地进行计算, 可把这个分解表示为矩阵形式. 我们要指出, 欧氏空间 $\mathbb{R}^n$ 的任意一个运动 g 可以唯一地表示为 $y = Ax + b$, 其中 y, x 为 $\mathbb{R}^n$ 中向量, $A \in O(n), b$ 为 $\mathbb{R}^n$ 中的一个固定向量, 即定义平移的向量. 空间 $\mathbb{R}^n$ 中的非齐次变换 g 可以唯一地相伴于下面的空间 $\mathbb{R}^{n+1}$ 中齐次变换

$$\widehat{g} = \left(\begin{array}{c|c} \multirow{3}{*}{A} & b^1 \\ & \vdots \\ & b^n \\ \hline 0 \cdots 0 & 1 \end{array}\right).$$

当这个算子 $\widehat{g}$ 作用在 $\mathbb{R}^{n+1}$ 时, 它在经过坐标为 $(0, \cdots, 0, 1)$ 的点而平行于坐标平面 $\mathbb{R}^n(x^1, \cdots, x^n)$ 的平面 $\mathbb{R}^n$ 上 (图 23) 生成了线性非齐次变换; 这个变换与空间 $\mathbb{R}^n$ 中的算子 g 的作用相同. 特别, 由算子 $\widehat{g}$ 的显式表示, 显然得到结论: 空间 $\mathbb{R}^{n+1}$ 的运动群分解为下面两个子群的半直积: 一个是旋转子群

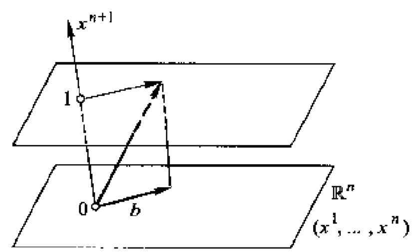


图 23

$$\left(\begin{array}{c|c} \multirow{3}{*}{A} & 0 \\ & \vdots \\ & 0 \\ \hline 0 \cdots 0 & 1 \end{array}\right),$$

另一个是平移子群

$$\left(\begin{array}{c|c} \multirow{3}{*}{O} & b^1 \\ & \vdots \\ & b^n \\ \hline 0 \cdots 0 & 1 \end{array}\right).$$

现在回到某个格 R 的变换群. 因为 $G_3(R) \subset G_3$, 则任意元 $g \in G_3(R)$ 有唯一的表示 $g = T\alpha$, 其中 T 为平移,α 为旋转 (一般说来,$T\alpha \neq \alpha T$). 我们强调指出一个重

要的情形: 变换 T 和 α 不必属于群 $G_3(R)$; 特别, 变换 T 完全不必是晶体平移群的元素 (但是复合算子 $T\alpha$ 已经是格把自己变到自己的变换).

群 $O(3)$(所有三维正交矩阵的群) 由两个连通分支组成, 即: 一个分支由行列式为 1 的矩阵组成 (正常旋转子群 $SO(3)$), 而第二个分支 (非正常旋转) 由行列式为 -1 的矩阵组成.$\mathbb{R}^3$ 中所有平行移动的群连通 (它由三个参数描述, 从而无异于三维空间自身). 同时, 群 $G_3(R)$ 为离散: 群 $G_3(R)$ 的离散性意味着在群 $G_3(R)$ 中不存在能随意靠近群的单位元的变换 (当然不包括恒同映射自己).

因此, 任意 $g \in G_3(R)$ 有形式 $g = T\alpha$, 其中变换 T 和 α 由 g 唯一地决定, 即由等式 $T\alpha = T'\alpha'$ 得到 $T = T', \alpha = \alpha'$ (参看前面的证明). 但是, 群 $G_3(R)$ 不能分解为它的两个子群的直积, 这是因为, 一般地 $T\alpha \neq \alpha T$.

定义 20.9. $O(3)$ 中所有那样的元 $\alpha \in O(3)$, 存在某个平移 T 使得变换 $T\alpha$ 属于群 $G_3(R)$ 的集合, 被称做晶体 (或格 R) 的点群.

换句话说, 旋转 α 属于晶体 R 的点群当且仅当存在这样一个平移 T(不必属于晶体平移群 $T_3(R)$) 使得复合变换 $T\alpha$ 属于晶体的空间群 $G_3(R)$. 我们用 $S_3(R)$ 表示这个点群. 也常常称此群为晶体 (格) 的对称群,而称它的元素为晶体 (格) 的对称运算.

命题 20.3. 变换的集合 $S_3(R)$ 构成通常代数意义下的群.

证明 因为群 G_3 的子群 T_3(平移群) 为其正规子群, 故可考虑由群 G_3 到商群 G_3/T_3 的投射 π. 这时群 $G_3(R)$ 变到了 G_3/T_3 的子群 $\pi G_3(R)$. 由早先证明过的知道, 商群 G_3/T_3 同构于群 $O(3)$, 因此 $\pi G_3(R)$ 是群 $O(3)$ 的子群. 这个子群的元素也仅仅是那些旋转使得存在元 $T \in T_3$, 有复合 $T\alpha$ 属于 $G_3(R)$. 因为与集合 $S_3(R)$ 的定义一致, 故而证明了集合 $S_3(R)$ 与 $\pi G_3(R)$ 重合, 特别, 它是一个群. 命题得证. □

注 因为 $\pi G_3(R)$ 同构于商群 $G_3(R)/(G_3(R) \cap \mathrm{Ker}\ \pi)$, 其中 $\pi : G_3 \to O(3), g = T\alpha \mapsto \alpha$, 还因为 $\mathrm{Ker}\ \pi = T_3$[①], 故最后有 $\pi G_3(R) \simeq S_3(R) \simeq G_3(R)/(G_3(R) \cap \mathrm{Ker}\ \pi)$. 另一方面, 交 $G_3(R) \cap T_3$ 与群 $T_3(R)$ 相同, 即晶体点群 $S_3(R)$ 同构于商群 $G_3(R)/T_3(R)$.

有时显式的算出晶体点群中的乘积运算和算出对应于两个对称的复合的平移是有用的. 设 $\alpha_1, \alpha_2 \in S_3(R)$. 考虑 $\alpha_3 = \alpha_1\alpha_2 \in S_3(R)$. 由 $S_3(R)$ 的定义, 存在这样的 T_1 和 T_2, 使 $T_1\alpha_1 \in G_3(R), T_2\alpha_2 \in G_3(R)$; 于是 $(T_1\alpha_1)(T_2\alpha_2) \in G_3(R)$. 然而 $T_1\alpha_1T_2\alpha_2 = T_1\alpha_1T_2\alpha_1^{-1} \cdot \alpha_1\alpha_2$, 因此作为对应于两个对称的复合的平移, 可以选取 $T_1\alpha_1T_2\alpha_1^{-1}$. 或者以显式的方式, 假定平移 T_1, T_2 由向量 x_1, x_2 定义, 而旋转 α_1, α_2 分别对应矩阵 A_1, A_2. 那么, 如果 r 为空间 $\mathbb{R}^3$ 中的向量, 则 $(T_1\alpha_1)r = A_1r + x_1, (T_2\alpha_2)r = A_2r + x_2$(先进行旋转而后平移). 于是 $(T_1\alpha_1)(T_2\alpha_2)r = (T_1\alpha_1T_2\alpha_2)r = (A_1A_2)r + (A_1x_2 + x_1)$. 因此, 由矩阵 A_1A_2 定义的旋转 $\alpha_3 = \alpha_1\alpha_2$ 转为变换 $g = T_3\alpha_3$, 其中 T_3 由 $x_3 = A_1x_2 + x_1$ 给出.

① 对同态 π 以 $\mathrm{Ker}\ \pi$ 代表群中使 $\pi(g) = 1$ 的元素的集合.

现在来给出平面 (二维) 格的例子, 这时存在 $g \in G_2(R)$, 其可分解为复合 $g = T\alpha$, 使得 $T \notin G_2(R)$ 而 $\alpha \notin G_2(R)$. 图 24 展示了一个格. 显然在直线 l 中的反射 $\alpha \in O(2)$ 不能保持格 R 不变; 另外, 按向量 β 的平移 (我们注意到这个平移不是格的基本平移) 也不能保持格 R 不变, 即这两个变换 T 和 α 不属于群 $G_2(R)$. 但是变换 $g = T\alpha$ 显然将格 R 变到自己. 这个运算 (变换)$g = T\alpha$ 被称做滑移反射. 我们再次强调指出, 晶体 (格) 点群中的元素 (变换), 一般地, 并不变晶体 (格) 到自己, 即晶体点群不是晶体运动群的子群. 点群在晶体结构理论中有重大的意义, 把它叫做格的对称群也并非毫无目的. 事实上, 因为它自身不仅包含有格的 "真正" 的对称, 并且有这样的变换, 它只在进行某个平移, 而且这个平移还不是个晶体运动后, 它才把格变到自己. 很清楚, 展示在图 24 中的格具有滑移反射的对称性. 除了具这种类型对称的晶体外, 常常遇到的是具有螺旋 (或轴向螺旋) 的晶体. 这种对称是旋转 $\alpha \in O(3)$ 和沿这个旋转轴的平移的复合. 建议读者去构造一个三维格的例子, 使它具有轴向螺旋的对称性.

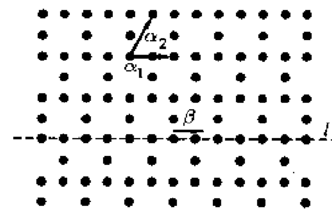


图 24

在晶体结构理论中还常常考察一个与每个格自然相联系的变换群.

定义 20.10. 群 $G_3(R)$ 中所有保持格不变且使旋转中心 O 固定不动的所有变换形成的子群被称为格 R 的*稳定群*$H_3(R)$.

我们记住, 点 O 被假定是不动的. 显见, 由于任意质点 O 不动的格的变换是正交变换, 即绕点 O 的旋转, 而且这个变换的行列式可以为 $+1$ 和 -1, 故 $H_3(R) = G_3(R) \cap O(3)$. 一般地说, 群 $H_3(R)$ 不是群 $G_3(R)$ 对于正交子群 $T_3(R)$ 的商群, 即 $H_3(R) \not\simeq G_3(R)/T_3(R) \simeq S_3(R)$. 参看图 24 的例子.

命题 20.4. 群 $H_3(R)$(稳定群),$S_3(R)$(点群或对称群) 和 $G_3(R)$(格的运动群) 在被考虑为群 G 的子群时, 它们之间有相关联的关系 $H_3(R) = S_3(R) \cap G_3(R)$.

证明 首先证明 $H_3(R) \subset S_3(R) \cap G_3(R)$. 设 $\alpha \in H_3(R)$. 特别, 当 α 为旋转 (从而属于 $G_3(R)$), 故可令 $g = T\alpha$, 其中 $T = E$ 为恒同变换 (按零向量位移), 即 $g = \alpha = E\alpha$; 因此, 由 $S_3(R)$ 的定义我们得到 $\alpha \in S_3(R)$. 反之, 我们证明 $H_3(R) \supset S_3(R) \cap G_3(R)$. 设 $\alpha \in S_3(R)$ 且 $\alpha \in G_3(R)$. 这意味着 α 保持格 R 不变, 且除此外还是个旋转, 即 $\alpha \in O(3) \cap G_3(R) = H_3(R)$.$\alpha$ 含在某个分解 $g = T\alpha$ 的这个事实在这种情形下是不存在的. 命题证完. □

注 如果格 R 为布喇菲格, 则群 $S_3(R)$ 和 $H_3(R)$ 相同. 这立即可由布喇菲格的定义得到 (参看前面).

远不是正交群 $O(3)$ 的每个子群都可以是某个格 R 的点群 (即对称群). 原来, 格的平移不变性在群 $S_3(R), G_3(R)$ 和 $H_3(R)$ 上加上了非常严格的限制. 我们以 $H_3(R)_{(0)}$ 表示群 $H_3(R)$ 的稳定子群, 它只由正常旋转组成, 即 $H_3(R)_{(0)} = SO(3) \cap G_3(R)$; 子群 $H_3(R)_{(0)}$ 中的每个旋转的行列式为 $+1$, 并使点 O 不动. 在晶体结构论

中下面的定理有重要的意义.

定理 20.1. 设 R 为平移不变格. 于是群 $H_3(R)_{(0)}$ 由有限个变换组成, 其中每一个是绕某轴的旋转, 此轴通过点 O, 旋转角为 $\pi/3$ 或 $\pi/2$ 的倍数.

证明 先考虑格 R 为布喇菲格的特殊情形, 即它的基本胞腔的所有原子均可由对点 O 施行格的基本平移而由 O 得到. 设 $\Phi\in H_3(R)_{(0)}$ 为正常旋转; 那么像由 §4 中所了解的, 变换 Φ 为绕某轴 l 旋转一个角 φ 的转动 (我们记得曾相对于点 O 计算过 $H_3(R)_{(0)}$, 而这个点 O 是格的一个原子). 设 Π 为垂直于 l 并通过点 O 的平面. 布喇非格由向量 $\alpha=n_1\alpha_1+n_2\alpha_2+n_3\alpha_3$ 所确定的点组成. 我们把格 R 的所有原子平行于直线 l 地投射到平面 Π 上, 并考虑位于点 O 的整个邻近的所有点, 但不包括 O 点本身. 我们固定其中的一个点 A_1 (多少可以表示它们位于离 O 有相等的距离. 参看图 25). 因为格 R 相对点 O 为对称, 故而跟随每个点 $B_1\in R$ 存在关于点 O 的反向点 B_1', 它仍属于 R. 当绕 l 旋转角 φ 时, 点 B_1 变到格的点 B_2 (因为变换 Φ 保持格 R 不变), 即投影 OA_1 变到投影 OA_2, 它们构成了角 φ. 因为向量 OB_1 和 OB_2 属于此格, 则它们的差向量 B_1B_2 也属于格 R. 因为向量 B_1B_2 平行于平面 Π, 因此在平移之后, 它确定了平面 Π 中向量 OA, 其端点 A 也属于此格. 由此得到向量 A_1A_2 的长度不小于向量 OA_1 的长 ($|OA_1|=|OA_2|$), 这是因为点 A_1 和 A_2 位于离点 O 的距离尽可能的小. 因此在三角形 OA_1A_2 中边 A_1A_2 不小于 $|OA_1|=|OA_2|$, 即角 φ 不小于 $\pi/3$. 进行一系列的变换 Φ, 我们在平面 Π 中得到正多边形, 其顶点为 $A_1,A_2,\cdots,A_m(A_{m+1}=A_1)$, 又因为 $\varphi\geqslant\pi/3$, 故 $1\leqslant m\leqslant 6$. 但是由于格 R 的对称性, 多边形关于 O 也是对称的. 因此,m 只可能取下列值:2,4,6. 因此角 φ 只可能等于后面的数:$k\pi,k\pi/2,k\pi/3$, 即 $\varphi=\pi/3,2\pi/3,\pi$. 于是对于布喇非格定理得证.

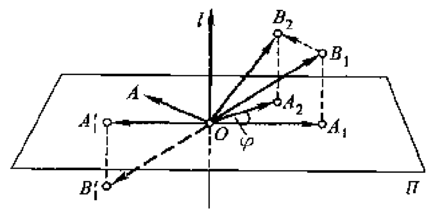


图 25

为了给出在一般情形的定理的证明, 还需证明, 在任一个格中都包含了一个布喇非子格, 它在旋转角 φ 时转换到自身. 但是我们将不在这里证明这个论断, 而宁愿给出另外一个证明, 它展现了一个重要的思想, 而这个思想在研究晶体结构时起了极大的作用. 我们先考虑平面格的情形. 以 α_1 和 α_2 代表格 R 的基本平移向量. 像以前那样,Φ 为绕点 O 的旋转, 旋转角为 φ 并将格变到自己. 考虑向量 α_1 并进行旋转 Φ 的迭代. 这样我们得到一系列向量 $\alpha_1,\Phi\alpha_1,\Phi^2\alpha_1,\cdots$. 因为基本格只包含了有限个原子, 则在进行有限次旋转 Φ后, 点应该回到最早的位置, 即 $\Phi^n\alpha_1=\alpha_1,n$ 为某个整数, 于是角 φ 有形式 $2m\pi/n,m,n$ 为某两个整数 (图 26). 因为向量 α_1 为平移向量, 并且因为平移子群为格的运动群的正规子群, 那么所有向量 $\alpha_1,\Phi\alpha_1,\Phi^2_{\alpha_1},\cdots$ 都是格的平移. 考虑三个向量 $\alpha_1,\Phi\alpha_1$ 和 $\Phi^2\alpha_1$. 如果向量 α_1 和 $\Phi^2\alpha_1$ 线性相关, 那么 Φ 便旋转了 $\pi/2$ 的倍数的角. 考虑 $\alpha_1,\Phi^2\alpha_1$ 为基底的情形. 这时为了使得格的运动

群为离散, 其充分必要条件是向量 $\Phi\alpha_1$ 按向量 $\alpha_1, \Phi^2\alpha_1$ 分解时具有有理的坐标. 事实上, 如果 $\Phi\alpha_1$ 在基 $\alpha_1, \Phi^2\alpha_1$ 下具无理数坐标, 那么, 进行一系列按向量 $\Phi\alpha_1$ 的平移, 在格的基本胞腔中便包含了无穷多个 (原子) 点, 这与格的定义矛盾. 下面我们要得到 $\Phi\alpha_1$ 在基 $\alpha_1, \Phi^2\alpha_1$ 下的坐标的解析表达式. 由图 27 我们得到了表达式

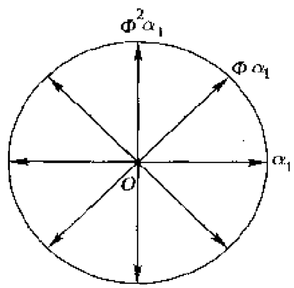


图 26

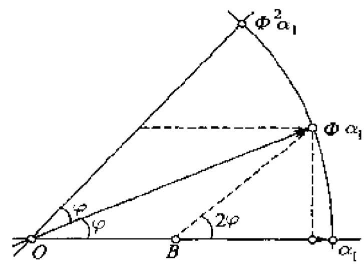


图 27

$$\Phi\alpha_1 = \frac{OB}{|\alpha_1|}(\alpha_1 + \Phi^2\alpha_1).$$

因此, $\dfrac{OB}{|\alpha_1|}$ 为有理数. 从同一图中清楚看出,

$$OB = |\alpha_1|\cos\ \varphi - OB\cos\ 2\varphi.$$

(这里的 φ 代表的角是 Φ 的旋转角.) 从而

$$\frac{OB}{|\alpha_1|} = \frac{1}{2\cos\ \varphi}.$$

那么向量 $\Phi\alpha_1$ 坐标的有理性等价于 $\cos\ \varphi$ 的有理性. 由此我们重新得到 φ 只能取值 $\pi/3, 2\pi/3, \pi/2, \pi$.

现在考虑在三维空间中任意使格不变的平移. 设 Φ 为格的一个任意旋转, 它绕点 O 且旋转角为 φ, 并使格变到自己. 考虑进行这个旋转所绕的轴 l 以及与轴 l 正交并经过 O 的平面 Π, 还有基本平移向量之一的向量 α_1(我们现在对其他两个基本平移向量暂时搁置一边). 当向量 α_1 落在平面 Π 中的情形, 则可逐字逐句地重复上面的讨论. 当 α_1 不在平面 Π 中时, 它在 Φ 作用下沿顶点为 O, 轴为直线 l 的锥面运动. 将此锥投射到平面 Π 上, 我们便能又重复前面关于向量在平面 Π 上投影的讨论. 于是定理彻底证好. □

现在考察平面格 R 和群 $H_2(R)_{(0)}$, 上面所证的定理允许我们对任意 (平移不变的) 平面格的群 $H_2(R)_{(0)}$ 的组成进行完全的描述. 现在我们列出五个群使任意群 $H_2(R)_{(0)}$ 等同于这五个群中的一个.

定理 20.2. (群 $H_2(R)_{(0)}$ 的分类定理). 设 $C_n(n = 1, 2, 3, 4, 6)$ 表示由下面形式的 n 个元素组成的群:

$$\begin{pmatrix} \cos\dfrac{2\pi k}{n} & \sin\dfrac{2\pi k}{n} \\ -\sin\dfrac{2\pi k}{n} & \cos\dfrac{2\pi k}{n} \end{pmatrix}, \quad 0 \leqslant k \leqslant n-1,$$

即群 C_n 由绕点 O 的, 角为 $2\pi k/n$ 的旋转组成. 于是对任意平面格 R 而言, 它的群 $H_2(R)_{(0)}$ 与群 C_n 中的一个相同, 其中 $n=1,2,3,4,6$.

这立刻从上面的定理得到.

因此, 平面格只能有上面列出的正常对称的类型. 我们指出, 群 C_1 只由恒同变换组成. 如果我们现在进一步不仅列出平面格的正常稳定群, 还列出非正常的群, 即全部的群 $H_2(R)$, 于是为此利用平面上的反射改变方向 (例如在 x 轴内的反射). 把此反射和群 $H_2(R)_{(0)}$ 中元组合, 从而考虑新群 $H_2(R)=H_2(R)_{(0)}\cup H_2(R)^*_{(0)}$, 我们显然得到了平面格的全部稳定律的完整清单

$$C_1, C_2, C_3, C_4, C_6; \quad D_1, D_2, D_3, D_4, D_6,$$

其中 D_i 是 C_i 和一个二阶循环群的半直积, 这个循环群是由保持格的反射生成. 因此 $C_i\subset D_i, i=1,2,3,4,6$, 且 $D_i/C_i\simeq\mathbb{Z}_2$; 特别 $D_1\simeq C_2$.

容易给出证明: 这些群中每一个对应于在这个群作用下不变的平面格. 我们将此作为习题留给读者.

我们注意到, 在所得到对称群的表格中缺少了 C_5. 因为平面平移不变的格在平面上定义了平面的不同图案, 许多装饰图案方面的专家对它们进行了数百年的研究. 特别, 在阿拉伯艺术中曾数次企图构造基于数字 5 的图案 (即同构于 C_5 的群 $H_2(R)_{(0)}$). 因为这些努力没有任何结果, 故而在阿拉伯的图案中引进了基于数 5 的所谓 "折衷变形", 它在某些地方破坏了装饰的对称性.

我们考虑任意平面的平移不变格. 我们已经描述过稳定群 $H_2(R)$ 所有可能的形式. 如何做出全部格的运动群 $G_2(R)$ 呢? 原来, 存在有正好 17 个互不同构的群, 它们给出了所有群 $G_2(R)$ 的完整清单. 我们在这里将不予证明. 这 17 个群中的每一个都对应着具这种群 $G_2(R)$ 的平面格. 有意思的是, 这 17 种图案也在古代就已显露出来了 (在主要的埃及装饰中).

我们现在转向三维格. 在这里的分类问题和完全描述所有类型的群 $H_3(R)$ 和 $G_3(R)$ 比起平面情形大大地复杂化. 我们因而不会详细地进行分类, 而是选择了最简单的问题: 如何建立三维空间中所有有限正常旋转群. 因为对于任意平移不变的三维格而言, 它的稳定群 (恰如对称群 $S_3(R)$) 是离散的 (从而是有限的) 群, 故而在做出群 $SO(3)$ 的全部有限子群的完全清单后, 我们则可估计出三维格的群 $H_3(R)_{(0)}$ 和 $S_3(R)_{(0)}$ 的 "上界" 清单.

我们先推出三维空间的有限旋转群的某些表格. 考虑一条通过点 O 的直线 l, 并假设 Π 为垂直于直线 l 且通过点 O. 考虑群 C_n(在平面 Π 中绕点 O 旋转 $2\pi/n$ 角的循环旋转群) 在平面 Π 中的作用. 很清楚这个群转换成了绕轴 l 的整个三维空

间的旋转群. 还是以字母 C_n 代表这些群. 这里 $n = 1,2,3,\cdots$, 而 C_1 只是由恒同变换构成. 除了群 C_n, 在平面 Π 上还有某个群 D_n 的作用. 平面 Π 相对于 Π 中某条直线 q 的反射, 在三维空间中可以实现为绕这条直线 q, 旋转角 π 的旋转. 因此这个平面 Π 的非正常旋转可以补足为三维空间中的正常旋转. 我们记如此产生的群为 D'_n. 这样一来, 群 D'_n 便由下面的变换组成: 所有由了群 C_n 来的变换, 除此之外还有绕平面 Π 上 n 根轴的所有三维空间旋转角 π 的旋转, 这 n 条轴相邻间的夹角等于 $\frac{2\pi}{2n} = \frac{\pi}{n}$. 应该指出, 与群 C_2 相同的群 D'_1 只由两个元素组成: 恒同变换和绕平面 Π 中唯一的直线的半旋转, 故而这两个群同构. 因此, 如果我们想要列出不同 (不同构的) 群的表格, 那么应该把 D'_1 去掉. 这样, 得到了下面的表列:$C_n, n = 1,2,3,\cdots; D'_n, n = 2,3,4,\cdots$.

除了这两个离散群的无穷序列外, 在三维空间中还存在一些更加奇异的变换群. 事实上, 我们考虑在三维空间中的五种正多面体 (立方体, 八面体, 二十面体, 四面体, 十二面体). 它们中的每一个都关联着一个正常旋转的有限群, 它把这个多面体变到自己. 以这种方法我们得到了五个有限群. 但是在这五个群中有相重合的. 事实上, 只有三个群是不同的: 四面体群, 立方体群和十二面体群. 我们更仔细地考察一下这个问题. 在立方体中内接一个球面, 在球面中内接一个八面体, 使得八面体的顶点重合于那些立方体边缘和上述球面的切点 (图 28). 可清楚看出, 所有把立方体变到自己的旋转也必保持八面体不变. 反之也对. 因此, 立方体和八面体的对称群重合 (即同构). 完全一样地可建立十二面体和二十面体的对称 (正常旋转) 群相同. 我们分别以 T, W, P 表示四面体, 立方体 (和八面体), 十二面体 (和二十面体) 的正常旋转群. 我们可以算出 (将其作为习题留给读者) 这几个群的阶分别为 12,24,60. 如果多面体的全部旋转群 (即包括了非正常旋转) 所得到的群 $\widetilde{T}, \widetilde{W}, \widetilde{P}$ 的阶分别为 24,48,120. 我们发现, 我们所表示了的群已完全列出了所有正常的离散旋转群.

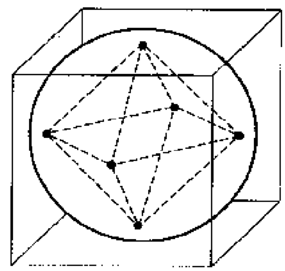


图 28

定理 20.3. 三维空间中有限正常旋转群的完整清单为

$$C_n \ (n = 1,2,\cdots),$$

$$D'_n \ (n = 2,3,\cdots),$$

$$T, W, P,$$

其中 C_n 为 (循环) 群, 它由重复进行角为 α 的绕某轴 l 的旋转构成, 其中 α 为 $2\pi/n, n$ 为整数;D'_n 是同样的旋转连同关于 n 条轴的反射, 这些轴是在垂直于 l 的平面中, 相互之间构成的角为 $\alpha/2; T, W, P$ 分别为保持正四面体, 立方体 (或八面体), 和十二面体 (或二十面体) 不变的变换群.

证明 首先我们已熟知, 在三维空间中非恒同变换的任意正常旋转为绕某条轴的旋转 (见 §4.3). 因此每个这样的旋转保持在单位球面上恰好两个对径的点不动, 这是旋转轴与此球面的交点. 我们称这两个点为旋转的极点.

现在设任给了一个有限群 Γ, 它为 N 阶正常旋转群. 于是其中正好有 $N-1$ 个非恒等变换的变换. 考虑它的极点. 称 Γ 中保持极点 p 不动的变换个数 ν 为极点 p 的重数. 显然, 这 ν 个变换是绕对应轴以角 $2\pi/\nu$ 的旋转的迭代. 它是 Γ 中 ν 阶的循环子群 Γ_p.Γ_p 中的非恒同的变换的个数为 $\nu-1$.

现在考虑关于群 Γ 的 p 点的轨道. 我们将证明, 重数为 ν 的极点 p 的轨道的元素个数正好等于 N/ν. 事实上, 我们注意到, 轨道上的每个元素都是具有相同重数 ν 的极点. 这个论断由后面的简单讨论得出: 设 Γ 中的变换 A 把 p 变到点 q, 而 $B\in\Gamma_p$, 即 B 使 p 保持不动. 于是变换 ABA^{-1} 把 q 变到 q. 反之, 如 T 为某个变换, 使 q 变到自己, 那么变换 $ATA^{-1}=B$ 将 p 点变到 p; 因此变换 T 具有形式 $A^{-1}BA$. 那么, 在群 Γ_p 的元素和保持点 q 不动的变换 T 之间建立了一一对应的关系, 从而证明了点 p 和 q 的重数相等. 以 $S_1, S_2, \cdots, S_\nu$ 表示保持点 p 不变的不同变换 (包括恒同变换). 我们假设在点 p 的轨道中恰好有 n 个点 (包括点 p). 以 $q_1,\cdots,q_n$ 记这些点. 现设 L_i 为变换群 Γ 中的那个元, 它将点 p 变到 q_i $(i=1,2,\cdots,n)$. 于是下面的这些不同变换穷竭了群 $\Gamma: L_1S_1, L_1S_2,\cdots,L_1S_\nu, L_2S_1, L_2S_2,\cdots,L_2S_\nu,\cdots,L_nS_1,L_nS_2,\cdots,L_nS_\nu$.

因此, $N=n\nu$, 这反过来表明这个重数为 N 的因子. 由此, 以 n_c 表示在轨道中的极点数, 它包含了最初的极点, 而以 ν_c 表示它们中每一个的重数时我们便得到 $N=n_c\nu_c$.

现在来计算偶对 (S,p) 的数目, 其中 p 为极点, 而 S 为群 Γ 的非恒同变换, 其中 S 使此极点不动. 一方面因为群 Γ 中 $N-1$ 个非恒同变换每一个正好有两个极点, 故此数目为 $2(N-1)$. 另一方面, 对每个极点 p 具有 ν_p-1 个非恒同变换保持这个极点不变 (此处 ν_p 为极点的重数). 于是这种偶对的数目等于 $\sum(\nu_p-1)$, 这里的 p 通过所有的极点. 由此得到

$$2(N-1)=\sum_c(\nu_c-1)n_c, \tag{1}$$

这里右端是对群 Γ 的所有轨道求和, 这些轨道包含了极点. 因为 $N=n_c\nu_c$, 于是以 N 除等式 (1) 便得到关系

$$2-\frac{2}{N}=\sum_c\left(1-\frac{1}{\nu_c}\right). \tag{2}$$

我们来研究等式 (2). 如果群 Γ 只由恒同变换组成, 则 $N=1$, 无极点. 现设 $N\geqslant 2$. 于是等式 (2) 的左端显然不小于 1, 但小于 2. 由此容易得知在等式的右端不能少于两项. 所以至少有两条轨道. 此时轨道的数目不能大于 3, 因为否则的话右端的和数将会不小于 2. 这样一来, 存在有两条或者三条轨道. 假设有两条轨道. 这时

等式 (2) 有形式 $2 = \frac{N}{\nu_1} + \frac{N}{\nu_2}$. 但两个整正数之和等于 2 只有它们每一个都等于 1. 因此 $n_1 = 1 = n_2$, 即 $\nu_1 = \nu_2 = N$. 这表明这两条轨道中每一条正好由一个重数为 N 的极点构成. 我们得到了绕 (竖直) 轴旋转的 N 阶循环群.

现在假设轨道的数目为 3. 于是

$$\frac{1}{\nu_1} + \frac{1}{\nu_2} + \frac{1}{\nu_3} = 1 + \frac{2}{N}. \tag{3}$$

我们把重数按其递增次序排列:$\nu_1 \leqslant \nu_2 \leqslant \nu_3$. 所有的这三个数 ν_1, ν_2, ν_3 不能都大于 2, 否则我们就会得到 $\frac{1}{\nu_1} + \frac{1}{\nu_2} + \frac{1}{\nu_3} \leqslant \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1$, 它与等式 (3) 矛盾. 因此 $\nu_1 = 2$. 由此,$\frac{1}{\nu_2} + \frac{1}{\nu_3} = \frac{1}{2} + \frac{2}{N}$. 这两个数 ν_2, ν_3 不能都 $\geqslant 4$, 因为此时在左端我们会得到不超过 1/2 的和数. 所以 $\nu_2 = 2$ 或 3. 考虑第一种情形 $\nu_1 = \nu_2 = 2, N = 2\nu_3$. 令 $\nu_3 = n$. 我们得到了两个重数为 2 的极点类, 每一类由 n 个极点组成, 以及另一个极点类, 它由两个重数为 n 的极点组成. 显然, 所得到的群与群 D'_n 相同 (参看前面).

现再考虑 $\nu_1 = 2, \nu_2 = 3, \frac{1}{\nu_3} = \frac{1}{6} + \frac{2}{N}$ 的情形. 因为 $\nu_3 \geqslant \nu_2 = 3$, 故有下面的三种可能性: $\nu_3 = 3, N = 12; \nu_3 = 4, N = 24; \nu_3 = 5, N = 60$, 我们分别以 T, W, P 代表它们. 就像我们现在看到的, 这些记号不是随意的, 它们是相应的正多面体的群的记号. 事实上, 我们来考察 T 的情形. 这时我们有两个类, 每个类由 4 个三阶极点 (即重数 3) 构成. 可清楚看出, 一个类的极点应该是正四面体的顶点, 而另一个类的极点为其对径点. 于是得到了四面体的正常旋转群. 6 个重数为 2 的等价极点是由它的中心 O 投射它的边的 6 个中点到球面的投影, 其中的这个球面内接于该四面体. 现在考虑 W 的情形. 这时, 一个类由 6 个四阶极点组成, 它们都是正八面体的顶点; 因此我们得到八面体的正常旋转群; 还有一个类由 8 个三阶极点组成 (对应于它的面的中心); 还有一个类由重数为 2 的 12 个极点组成 (对应于它的边的中点). 最后, 考虑 P 的情形. 这时我们有这样一些类: 由重数 5 的 12 个极点组成的类 (它们是正二十面体的顶点); 由 20 个重数为 3 的极点组成, 它们对应于 20 个面的中心点; 由 30 个重数为 2 的点组成, 它们是这个多面体的 30 条边的中点.

因此, 这些群的描述和定理的证明已完成. □

现在回到对三维格的稳定群的描述. 先进行对某些新的群的考察. 设 B 为相对于点 O 的三维空间的反射, 即变换 B 将点 X 变到点 $-X$. 显然这个正交变换改变了 $\mathbb{R}^3$ 的方向. 另外,B 与任意旋转可交换, 即 B 为群 $O(3)$ 的中心的一个生成元; 用矩阵记号即 $B = -1$, 其中 1 是单位矩阵. 于是 $SO(3)$ 中每个子群 H 生成了群 $O(3)$ 中的子群 $\widetilde{H}$, 它由 H 加上所有形如 Bg 的元构成, 其中 $g \in H$. 显然,$\widetilde{H}$ 包含了 H 作为它的指数 2 的子群.

我们还要描述一种把镜像反射包括在有限群的方法. 设 Γ 为在另一正常旋转群

Φ 中的指数为 2 的子群, 而只群 Φ 的一半元素 (记为 S) 属于 Γ, 而 Φ 的另一半元 (记为 S') 不属于 Γ. 现将 S' 中所有元素变换为 BS' 中元. 于是我们得一个新的包含 Γ 的群, 记其为 $\Phi\Gamma$, 同时作为这个群的另一半元素 (即不属于 Γ 的那些元素) 组成了镜像反射. 例如, 因为四面体群 T 在八面体群 W 中为指数 2 的子群, 故而有群 WT.

定理 20.4. 三维格的平移不变的全部稳定群的清单 (包括了 32 个群, 它们互不同构) 为: $C_1, C_2, C_3, C_4, C_6, \widetilde{C}_1, \widetilde{C}_2, \widetilde{C}_3, \widetilde{C}_4, \widetilde{C}_6; D_2', D_3'; D_4', D_6', \widetilde{D}_2', \widetilde{D}_3', \widetilde{D}_4', \widetilde{D}_6'; C_2C_1, C_4C_2, C_6C_3; D_4'D_2', D_6'D_3'; D_2'C_2, D_3'C_3, D_4'C_4, D_6'C_6; T, W, \widetilde{T}, \widetilde{W}, WT$. 这些群中的每一个都满足三维格的平移不变性.

证明 结合前面已证过的描述所有正常旋转的有限群的定理, 并进而结合描述包括镜像反射的过程, 又最后注意到根据所允许的只有重数为 2,3,4,6 的极点的那些情形, 便能给出证明. 我们将不给出这些类比的所有细节. □

现在来考虑与具体的张量场有关的晶体群. 我们局限于一个简单的例子: 挑出晶体的一个宏观的性质, 就像导电性描述了电场电压向量 ε 和电流密度向量 j 之间关联那样. 这种关联有下面的形式

$$j_k = \sigma_k^s \varepsilon_s, \quad 其中\ j = (j_k); \varepsilon = (\varepsilon_s),$$

而 (σ_k^s)为物质的导电率张量. 如果物质是各向同性的, 则有 $\sigma_k^s = \sigma\delta_k^s$. 其中 σ 为标量, 即在此最简单情形中导电率只由一个标量 σ 给出. 在一般情形,(σ_k^s) 为一般形式的张量.

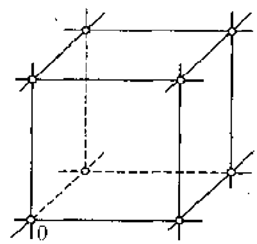


图 29

考虑晶体的导电率张量, 它由 $\mathbb{R}^3$ 中的立方体状的格描述, 即为立方体晶体 (图 29). 我们认为晶体是理想的, 从而格 R 充满了整个空间. 显然, 这个晶体的对称群特别地含有下面三个正交变换

$$\alpha_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \quad \alpha_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}; \quad \alpha_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix},$$

即 α_1 为绕 z 轴的角为 $\pi/2$ 的旋转,α_2 为绕 y 轴的角为 $\pi/2$ 的旋转,α_3 为绕 x 轴的角为 $\pi/2$ 的旋转.

因为格 R 变到自己, 故而这三个对称运算保持了张量 (σ_k^s) 不变. 我们把这种情形写成分析形式. 以 A 表矩阵 (σ_k^s); 于是 $A_i' = \alpha_i A \alpha_i^{-1} = A$ 对 $i = 1,2,3$ 都成立. 计算矩阵 A_1', 得到

$$A_1' = \begin{pmatrix} \sigma_2^2 & -\sigma_1^2 & \sigma_3^2 \\ -\sigma_2^1 & \sigma_1^1 & -\sigma_3^1 \\ \sigma_2^3 & -\sigma_1^3 & \sigma_3^3 \end{pmatrix} = \begin{pmatrix} \sigma_1^1 & \sigma_2^1 & \sigma_3^1 \\ \sigma_1^2 & \sigma_2^2 & \sigma_3^2 \\ \sigma_1^3 & \sigma_2^3 & \sigma_3^3 \end{pmatrix} = A.$$

由此得出 $\sigma_1^1=\sigma_2^2$; 以相似的方法计算 A_2', A_3', 我们得到 $\sigma_1^1=\sigma_2^2=\sigma_3^3$. 群 $S_3(R)$ 还包含了三个变换

$$\beta_1=\begin{pmatrix}-1&0&0\\0&-1&0\\0&0&1\end{pmatrix},\quad \beta_2=\begin{pmatrix}1&0&0\\0&-1&0\\0&0&-1\end{pmatrix},\quad \beta_3=\begin{pmatrix}-1&0&0\\0&1&0\\0&0&-1\end{pmatrix},$$

于是 β_1 为绕 z 轴的角为 π 的旋转,β_2 为绕轴 x 的角为 π 的旋转,β_3 为绕 y 轴的角为 π 的旋转. 此时格又变到自己, 这给出了关系 $\widetilde{A}_i=\beta_iA\beta_i^{-1}=A, 1\leqslant i\leqslant 3$. 计算矩阵 $\widetilde{A}_1$ 给出了

$$\widetilde{A}_1=\begin{pmatrix}\sigma_1^1&\sigma_2^1&-\sigma_3^1\\\sigma_1^2&\sigma_2^2&-\sigma_3^2\\-\sigma_1^3&-\sigma_2^3&-\sigma_3^3\end{pmatrix}=\begin{pmatrix}\sigma_1^1&\sigma_2^1&\sigma_3^1\\\sigma_1^2&\sigma_2^2&\sigma_3^2\\\sigma_1^3&\sigma_2^3&\sigma_3^3\end{pmatrix}=A,$$

由此得出 $\sigma_3^1=\sigma_3^2=\sigma_1^3=\sigma_2^3=0$. 计算矩阵 $\widetilde{A}_2$ 和 $\widetilde{A}_3$. 类似地可得出 $\sigma_j^i=0$, 当 $i\neq j$, 即最终有 $A=(\sigma_k^s)=\sigma\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$, 即 $\sigma_k^s=\sigma\delta_k^s$, 其中 σ 为标量. 我们从而证明了下面的断言: 立方体晶体的导电率的迷向性类似于各向同性物质的导电率 (即不依赖于方向). 这个结论在物理上并不是显见的, 因为在立方体格的情形我们会预料在立方体边的方向上的导电性不同于对角线方向的导电性. 因此我们显示了对称群 $S_3(R)$ 性质的重要应用 (虽然是简单的), 这个性质能出人意料地降低张量 (σ_k^s) 中独立的分量数目.

§21. 伪欧氏空间的二阶张量和它们的特征值

1. 反称张量. 电磁场的不变量

我们考查重要张量的例子: 闵可夫斯基空间 $\mathbb{R}^4_{1,3}$ 中的二阶反称张量. 特别, 电磁场F_{ik} 是那种我们采用相应的术语“张量场”的张量.

定义 21.1. 张量场 F_{ik} 的不变量是指特征多项式

$$P(\lambda)=\det(F_{ik}-\lambda g_{ik}) \tag{1}$$

的系数, 其中 g_{ik} 为闵可夫斯基度量.

我们引进电场$\mathbf{E}$ 和磁场 $\mathbf{H}$, 其中

$$E_\alpha=F_{0\alpha},\quad \alpha=1,2,3;$$

$$-H^1=F_{23},\quad -H^2=F_{31},\quad -H^3=F_{12}. \tag{2}$$

那么, 矩阵 F_{ik} 因此具下面形式

$$F_{ik}=\begin{pmatrix}0 & E_1 & E_2 & E_3\\ -E_1 & 0 & -H^3 & H^2\\ -E_2 & H^3 & 0 & -H^1\\ -E_3 & -H^2 & H^1 & 0\end{pmatrix}. \tag{3}$$

引理 21.1. 张量 F_{ik} 的特征多项式为

$$P(\lambda)=-\lambda^4+(\mathbf{E}^2-\mathbf{H}^2)\lambda^2+\langle\mathbf{E},\mathbf{H}\rangle^2, \tag{4}$$

其中 $|\mathbf{E}|^2=E_1^2+E_2^2+E_3^2, |\mathbf{H}|^2=(H^1)^2+(H^2)^2+(H^3)^2, \langle\mathbf{E},\mathbf{H}\rangle=E_iH^i$.

证明 经过直接计算矩阵 $(F_{ik}-\lambda g_{ik})$ 的行列式就可得到, 其中 F_{ik} 有形式 (3). 借助于下面所证明的定理 21.1 可以容易得到另外一个证明, 这是个关于在伪欧氏空间中关于它的反称张量标准形的定理. □

注 虽然 F_{ik} 在所给点上是反称的, 但由于度量的非欧性, 一般说来它不能通过洛伦兹变换化为标准的块状形式.

将公式 (2) 转化为微分形式的表达方式得到

$$\begin{aligned}F&=\sum_{i<j}F_{ij}dx^i\wedge dx^j\\ &=E_\alpha dx^0\wedge dx^\alpha-H^1dx^2\wedge dx^3-H^2dx^3\wedge dx^1-H^3dx^1\wedge dx^2.\end{aligned} \tag{5}$$

向量 $\mathbf{E}$ 和 $\mathbf{H}$ 是相对于三维空间 (x^1,x^2,x^3) 的旋转而言的向量. 因为 F_{ik} 为张量, 那么在洛伦兹变换

$$x^1=\frac{x^{1'}+vt'}{\sqrt{1-\frac{v^2}{c^2}}};\quad x^2=x^{2'};\quad x^3=x^{3'};\quad t=\frac{t'+\frac{v}{c^2}x^{1'}}{\sqrt{1-\frac{v^2}{c^2}}}, \tag{6}$$

下, 其中 $x^0=ct, x^{0'}=ct'$, 向量 $\mathbf{E}$ 和 $\mathbf{H}$ 为

$$\begin{aligned}&E_1=E_1',\quad E_2=\frac{E_2'+\frac{v}{c}H^{3'}}{\sqrt{1-\frac{v^2}{c^2}}},\quad E_3=\frac{E_3'-\frac{v}{c}H^{2'}}{\sqrt{1-\frac{v^2}{c^2}}};\\ &H^1=H^{1'},\quad H^2=\frac{H^{2'}-\frac{v}{c}E_3'}{\sqrt{1-\frac{v^2}{c^2}}},\quad H^3=\frac{H^{3'}+\frac{v}{c}E_2'}{\sqrt{1-\frac{v^2}{c^2}}}.\end{aligned} \tag{7}$$

在给定点上张量 F_{ik} 构成了六维向量空间. 我们要解释在此空间中算子 * 是如何作用的.

引理 21.2. 成立公式

$$*F = -\sum_{\alpha=1}^{3} H^\alpha dx^0 \wedge dx^\alpha - E_1 dx^2 \wedge dx^3 - E_2 dx^3 \wedge dx^1 - E_3 dx^1 \wedge dx^2, \tag{8}$$

其中 F 的形式为

$$F = E_\alpha dx^0 \wedge dx^\alpha - H^1 dx^2 \wedge dx^3 - H^2 dx^3 \wedge dx^1 - H^3 dx^1 \wedge dx^2.$$

证明 算子 $*$ 作用如下:

$$(*F)_{ik} = \frac{1}{2}\varepsilon_{iklm} F^{lm}; \quad F^{lm} = g^{lp} g^{mq} F_{pq}. \tag{9}$$

这里的 g_{ij} 为闵可夫斯基度量,

$$g_{ij} = \begin{pmatrix} 1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & -1 \end{pmatrix};$$

张量 F^{ik} 有形式

$$F^{0\alpha} = -F_{0\alpha} = -E_\alpha; \quad F^{\alpha\beta} = F_{\alpha\beta}, \quad \alpha, \beta = 1, 2, 3. \tag{10}$$

由此立即得到了引理的断言. □

推论 作用于二阶反称张量的双重 $*$ 运算等于 -1,
即 $*(*F) = -F$.
令

$$(a + bi)F = aF + b * F \tag{11}$$

给出了反称 2 阶张量的空间 $\mathbb{R}^6$ 一个复线性空间结构.

由于上述推论, 这个定义是合理的:$i^2 F = * * F = -F$. 我们从而得到了三维的复空间 $\mathbb{C}^3$.

注 我们将认为 $\mathbf{E}$ 是 F 的实部,$\mathbf{H}$ 是它的虚部. 这与前述定义并不抵触. 事实上, 对 $\mathbf{E} + i\mathbf{H}$ 乘以 i 有

$$i(\mathbf{E} + i\mathbf{H}) = -\mathbf{H} + i\mathbf{E}.$$

由于有引理 21.2, 这完全是算子 $*$ 的作用. 因此, 令

$$z^\alpha = E_\alpha + iH^\alpha, \quad \alpha = 1, 2, 3, \tag{12}$$

便能引进了 $\mathbb{C}^3$ 中坐标 z^1, z^2, z^3.

算子 $*$ 对于群 $SO(1,3)$ 中变换不变. 于是最后的这个公式是我们的空间 $\mathbb{C}^3$ 的复线性变换. 进而容易看出, 它们保持了二次型 $\langle F,F\rangle$ 不变, 现在给出其定义如后. 为此首先注意到 $F\wedge F$ 和 $F\wedge *F$ 都是 (0,4) 型反称张量, 并且因为闵可夫斯基空间是四维的, 故这两个张量中每一个的所有 (非零) 分量都相等 (差一个符号). 直接计算给出

$$(F\wedge *F)_{0123}=\frac{1}{2}F_{ik}F^{ik};\quad (F\wedge F)_{0123}=-\frac{1}{2}\varepsilon^{ijkl}F_{ij}F_{kl}.$$

标量 $*(F\wedge F)$ 和 $*(F\wedge *F)$ 的任意线性组合仍是标量, 故而二次型

$$\begin{aligned}\langle F,F\rangle &= *(F\wedge *F+iF\wedge F)\\ &=-\frac{1}{2}(F_{ik}F^{ki}+i\varepsilon^{ijkl}F_{ij}F_{kl})\end{aligned}\tag{13}$$

对于群 $SO(1,3)$ 中变换不变. 用坐标 z^1,z^2,z^3 写出这个形式, 我们有

$$\langle F,F\rangle=-H^2+E^2+2i\langle\mathbf{E},\mathbf{H}\rangle=(\mathbf{E}+i\mathbf{H})^2=\sum_{\alpha=1}^{3}(z^\alpha)^2\tag{14}$$

(其中 $E^2=|\mathbf{E}|^2,H^2=|\mathbf{H}|^2$). 于是, 形式 (13) 为复三维向量 $\mathbf{E}+i\mathbf{H}$ 的标量式的平方.

空间 $\mathbb{C}^3$ 所保持这个标量平方不变的复线性变换, 类比于实的情形被称做复正交, 它构成了一个群, 我们以 $O(3,\mathbb{C})$ 记之; 不要把它与 $U(3)$ 混同.

因此, 在 2 阶反称张量的空间中所引进复结构便定义了洛伦兹群 $SO(1,3)$ 到 $O(3,\mathbb{C})$ 的一个同态.

而且量 $\mathrm{Re}\langle F,F\rangle=E^2-H^2,\dfrac{1}{2}\mathrm{Im}\langle F,F\rangle=\langle\mathbf{E},\mathbf{H}\rangle$ 为电磁场的不变量.

现在来分析用洛伦兹变换化 2 阶反称张量为标准形.

定理 21.1. 1) 设 $\langle F,F\rangle=E^2-H^2+2i\langle\mathbf{E},\mathbf{H}\rangle\neq 0$.

a) 如果 $\langle\mathbf{E},\mathbf{H}\rangle\neq 0$, 则洛伦兹变换可以把张量 F 变到这样的形式, 使得 $\mathbf{E}$ 和 $\mathbf{H}$ 平行且都不为零.

b) 如果 $\langle\mathbf{E},\mathbf{H}\rangle=0,E^2-H^2\neq 0$, 则可将张量 F 变成这样的形式, 使得当 $E^2-H^2>0$ 时 $\mathbf{E}\neq 0,\mathbf{H}=0$, 或者当 $E^2-H^2<0$ 时 $\mathbf{E}=0,\mathbf{H}\neq 0$.

张量 F 的标准形式在这两种情形下都为

$$F=\begin{pmatrix}0&E'&0&0\\-E'&0&0&0\\0&0&0&-H'\\0&0&H'&0\end{pmatrix},\tag{15}$$

其中 $E'^2-H'^2=E^2-H^2,E'H'=\langle\mathbf{E},\mathbf{H}\rangle$.

2) 设 $\langle F,F\rangle=0$, 即 $E^2-H^2=0,\langle\mathbf{E},\mathbf{H}\rangle=0$. 于是经任意洛伦兹变换向量 $\mathbf{E}$ 和 $\mathbf{H}$ 仍相互垂直并且有相同长度. 张量 F_{ik} 在这时可化为

$$F=\begin{pmatrix}0&E&0&0\\-E&0&-E&0\\0&E&0&0\\0&0&0&0\end{pmatrix}. \tag{16}$$

证明 设 $\langle F,F\rangle\neq0$. 令 $F=c\cdot n$, 其中 n 为 $\mathbb{C}^3$ 中的单位向量, 即 $c=\sqrt{\langle F,F\rangle}$ 且 $\langle n,n\rangle=1$, 用 $SO(3,\mathbb{C})$ 中元旋转 (这种旋转是洛伦兹变换) 把向量 n 可以变到其他任意一个向量, 特别变到实向量 n'. 我们表示 c 为形式 $E'+iH'$; 新的电场向量有形式

$$\mathbf{E}'=E'n', \tag{17}$$

而新的磁场向量的形式为

$$\mathbf{H}'=H'n', \tag{18}$$

因而 $c^2=E'^2-H'^2+2i\langle\mathbf{E}',\mathbf{H}'\rangle=\langle F,F\rangle=E^2-H^2+2i\langle\mathbf{E},\mathbf{H}\rangle$. 向量 $\mathbf{E}'$ 和 $\mathbf{H}'$ 平行. 我们选取 n' 为 z^1 轴的单位向量, 则我们得到了标准形 (15). 定理 21.1 论断 1) 的 a) 得证.

如果 $\langle\mathbf{E},\mathbf{H}\rangle=0$, 则 $\langle\mathbf{E}',\mathbf{H}'\rangle=\langle\mathbf{E},\mathbf{H}\rangle=0$. 因此或者 $E'=0$, 或 $H'=0$. 从而论断 1) 的 b) 得证: 如果 $E^2-H^2>0$, 则 $H'=0$; 如果 $E^2-H^2<0$, 则 $E'=0$.

最后我们证明 2).$SO(3)\subset SO(1,3)$ 中的实旋转可以使得向量 $\mathbf{E}$ 指向 z^1 轴的方向. 于是向量 $\mathbf{H}$ 位于 (z^2,z^3) 平面中. 在此平面中的旋转可以使它指向 z^3 轴的方向. 定理证完. □

2. 对称张量和特征值. 电磁场的能量 – 动量张量

设 T_{ik} 为闵可夫斯基空间中 (0,2) 型的对称张量, 而此空间的度量为 g_{ik}.

定义 21.2. 称方程

$$P(\lambda)=\det(T_{ik}-\lambda g_{ik})=0 \tag{19}$$

的解为**张量 T_{ik} 的特征值**.

我们记得, 这些特征值 λ 为算子 $T^i_k=g^{ij}T_{jk}$ 的特征值. 相应的特征向量由方程组

$$T_{ik}\xi^k=\lambda g_{ik}\xi^k,\quad P(\lambda)=0 \tag{20}$$

的解定义. 因为度量的伪欧氏性, 一般说来矩阵 T_{ik} 不能由洛伦兹变换化为对角形. 我们要详细地分析在二维闵可夫斯基空间 $\mathbb{R}^2_1$ 中对称张量的类型. 有下面定理.

定理 21.2. 1) 如果方程 $P(\lambda)=0$ 有两个不同的实根 $\lambda_0\neq\lambda_1$, 则矩阵 T_{ik}(由洛伦兹变换) 化为

$$\begin{pmatrix}\lambda_0 & 0\\ 0 & -\lambda_1\end{pmatrix}. \tag{21}$$

2) 如果方程 $P(\lambda)=0$ 有两个复共轭根 $\lambda_0,\lambda_1=\alpha\pm i\beta$, 则矩阵 T_{ik} 可化为

$$\begin{pmatrix}\alpha & \beta\\ \beta & -\alpha\end{pmatrix}. \tag{22}$$

3) 如果方程 $P(\lambda)=0$ 有两个相同的根 $\lambda_0=\lambda_1=\lambda$, 则在任意坐标下矩阵 T_{ik} 有形式

$$\begin{pmatrix}\lambda+\mu & -\mu\\ -\mu & -\lambda+\mu\end{pmatrix}, \tag{23}$$

这里的 μ 不是张量 T_{ik} 的不变量, 并且一般说来, 不能用洛伦兹变换化为 0.

证明 设特征值 $\lambda_0\neq\lambda_1$ (实或复) 不相同. 它们对应了两个线性无关的特征向量 ξ_0,ξ_1, 它们满足方程组 (20). 我们指出, 这些向量正交: 我们有 $\xi_0^iT_{ik}\xi_1^k=\lambda_0g_{ik}\xi_0^i\xi_1^k=\lambda_1g_{ik}\xi_0^i\xi_1^k$, 故而 $\langle\xi_0,\xi_1\rangle=g_{ik}\xi_0^i\xi_1^k=0$. 假设 $\lambda_0\neq\lambda_1$ 为实数. 那么其中一个特征向量是类时的, 而另一个是类空的. 将其选为新的坐标轴, 我们则得到了定理 21.2 的论断 1) 的证明.

假设 $\lambda_0,\lambda_1=\alpha\pm i\beta$, 其对应的特征向量也为复共轭, 即 ξ_0,ξ_1 分别为 $a+ib$ 和 $a-ib$. 对这两个向量以下面条件法化

$$\langle a+ib,a+ib\rangle=\langle a-ib,a-ib\rangle=2.$$

与正交条件 $\langle\xi_0,\xi_1\rangle=0$ 一起给出了

$$\langle a,b\rangle=0,\quad \langle a,a\rangle+\langle b,b\rangle=0,\quad \langle a,a\rangle-\langle b,b\rangle=2.$$

由此,$|a|^2=1,|b|^2=-1$. 在基 a,b 下张量 (T_{ik}) 有了形式 (22). 论断 2) 得证.

现在假设 $\lambda_0=\lambda_1=\lambda$. 我们假定特征值 λ 恰好对应于一个特征向量 ξ. (如果这样的向量有两个, 那么矩阵 T_{ik} 可转化为对角形 $\begin{pmatrix}\lambda & 0\\ 0 & -\lambda\end{pmatrix}$, 即为 (23) 中 $\mu=0$ 的情形.) 我们指出 ξ 是迷向向量. 事实上, 在相反的情形下,ξ 的正交补向量就会是特征子空间, 且不包含 ξ. 因此 $|\xi|^2=0$, 并由此得到

$$(\xi^0)^2-(\xi^1)^2=0\ \Rightarrow\xi^0=\xi^1$$

(如果 $\xi^0=-\xi^1$, 则还需进行反射). 方程组 (20) 因而可写为

$$T_{00}+T_{01}=\lambda,\quad T_{10}+T_{11}=-\lambda,$$

由此有 $T_{00}=\lambda+\mu,T_{11}=-\lambda+\mu,T_{01}=-\mu=T_{10}$. 定理得证. □

假设 F_{ik} 为在闵可夫斯基空间 $\mathbb{R}^4_1$ 中的反称张量. 由张量 F_{ik} 我们可以构造出对称张量 T_{ik}, 这时令

$$T_{ik} = \frac{1}{4\pi}\left(-g^{lm}F_{il}F_{km} + \frac{1}{4}F_{lm}F^{lm}g_{ik}\right). \tag{24}$$

如果 F_{ik} 为电磁场张量, 则称张量 T_{ik} 为场 F_{ik} 的能量 – 动量张量. 利用由公式 (2) 定义的三维向量 $\mathbf{E}, \mathbf{H}$, 我们得到

$$T_{00} = \frac{E^2+H^2}{8\pi}, \quad T_{0\alpha} = -\frac{1}{4\pi}[\mathbf{E}, \mathbf{H}]_\alpha,$$

$$T_{\alpha\beta} = \frac{1}{4\pi}\{-E_\alpha E_\beta - H_\alpha H_\beta + \frac{1}{2}\delta_{\alpha\beta}(E^2+H^2)\}, \quad \alpha, \beta = 1, 2, 3. \tag{25}$$

称 $W = T_{00}$ 为电磁场 F_{ik} 的能量密度; 称向量 $S_\alpha = -T_{0\alpha}$ 为坡印廷向量; 称三维张量 $T_{\alpha\beta}$ $(1\leqslant\alpha\leqslant3, 1\leqslant\beta\leqslant3)$ 为麦克斯韦应力张量.

我们来研究关于把由公式 (24) 定义的能量 – 动量张量化为标准形的问题. 根据定理 21.1, 我们化张量 F_{ik} 为标准形. 在 (15) 的情形, 有 $\mathbf{E} = (E, 0, 0), \mathbf{H} = (H, 0, 0)$, 那么张量 T_{ik} 便具有对角形

$$(T_{ik}) = \begin{pmatrix} +W & & & 0 \\ & -W & & \\ & & +W & \\ 0 & & & +W \end{pmatrix}, \quad W = \frac{1}{8\pi}(E^2+H^2). \tag{26}$$

这里所有特征值等于 $\pm W$. 而在情形 (16), 有

$$\mathbf{E} = (E, 0, 0), \quad \mathbf{H} = (0, 0, H), \quad E = H,$$

由公式 (25) 我们得到

$$(T_{ik}) = \begin{pmatrix} W & 0 & -W & 0 \\ 0 & 0 & 0 & 0 \\ -W & 0 & W & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad W = \frac{E^2}{4\pi} = \frac{H^2}{4\pi}. \tag{27}$$

这种情形对应于定理 21.2 的最后一个论断 (公式 (23)), $\mu = W$ 为电磁波的情形. 能量 – 动量张量此时不能化为对角形; 它所有的特征值为 0.

§22. 在映射下张量的行为

1. 具下指标的张量的一般限制运算

设给出从具坐标 $x^{1'}, \cdots, x^{m'}$ 的 m 维空间的区域到具坐标 $x^1, \cdots, x^n$ 的 n 维空间的区域的映射 F 为

$$x^i = x^i(x^{1'}, \cdots, x^{m'}), \quad i = 1, \cdots, n. \tag{1}$$

那么在空间 $(x^1, \cdots, x^n)$ 中 $(0,k)$ 型的每个张量 $T_{i_1\cdots i_k}$ 对应于"带撇"的空间中的张量 $(F^*T)_{i'_1\cdots i'_k}$:

$$(F^*T)_{i'_1\cdots i'_k}(x^{1'}, \cdots, x^{m'}) = \left[T_{i_1\cdots i_k}\frac{\partial x^{i_1}}{\partial x^{i'_1}}\cdots\frac{\partial x^{i_k}}{\partial x^{i'_k}}\right](x^i(x^{1'}, \cdots, x^{m'})). \tag{2}$$

我们把验证 F^*T 为 $(0,k)$ 型张量留给读者.

因此,$(0,k)$ 型的张量被映射的方向与空间之间映射的方向相反. 称运算 F^* 为张量的限制运算.

例 22.1. 设在 n 维空间中给定了度量 g_{ij} 和 m 维曲面

$$x^i = x^i(x^{1'}, \cdots, x^{m'}), \quad i = 1, \cdots, n. \tag{3}$$

于是应用限制运算我们得到在曲面 (3) 的度量 $g_{i'j'}$, 其中

$$g_{i'j'} = g_{ij}\frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^j}{\partial x^{j'}}, \quad i', j' = 1, \cdots, m.$$

这是曲面所在空间的度量 g_{ij} 在此曲面上诱导的度量 (参看 §7.3).

现在考察 $(0,k)$ 型反称张量 $T_{i_1\cdots i_k}$ 在 k 维曲面 $x^i = x^i(x^{1'}, \cdots, x^{k'})$ 上限制的情形, 其中这个曲面是在 n 维空间中的. 这个限制运算的显式表达式在我们的反称形式的积分论中有用.

定理 22.1. 限制于 k 维曲面 $x^i = x^i(x^{1'}, \cdots, x^{k'})$ 上的形式

$$\sum_{i_1<\cdots<i_k} T_{i_1\cdots i_k}dx^{i_1}\wedge\cdots\wedge dx^{i_k}$$

成立下面等式

$$\sum_{i_1<\cdots<i_k} T_{i_1\cdots i_k}dx^{i_1}\wedge\cdots\wedge dx^{i_k} = \sum_{i_1<\cdots<i_k} J^{i_1\cdots i_k}T_{i_1\cdots i_k}dx^{1'}\wedge\cdots\wedge dx^{k'}, \tag{4}$$

其中 $J^{i_1\cdots i_k}$ 为矩阵 $\left(\dfrac{\partial x^i}{\partial x^{i'}}\right)$ 的 k 阶子式, 由 $i_1, \cdots, i_k$ 列组成.

证明 根据定义 (2), 有

$$(F^*T)_{1'\cdots k'} = T_{i_1\cdots i_k}\frac{\partial x^{i_1}}{\partial x^{1'}}\cdots\frac{\partial x^{i_k}}{\partial x^{k'}}. \tag{5}$$

利用 $T_{i_1\cdots i_k}$ 的反称性我们可以将上式右端化为

$$\begin{aligned} T_{i_1\cdots i_k}\frac{\partial x^{i_1}}{\partial x^{1'}}\cdots\frac{\partial x^{i_k}}{\partial x^{k'}} &= \sum_{i_1<\cdots<i_k} T_{i_1\cdots i_k}\left(\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)\frac{\partial x^{i_{\sigma(1)}}}{\partial x^{1'}}\cdots\frac{\partial x^{i_{\sigma(k)}}}{\partial x^{k'}}\right) \\ &= \sum_{i_1<\cdots<i_k} T_{i_1\cdots i_k}J^{i_1\cdots i_k}, \end{aligned}$$

从而定理得证. □

2. 切空间的映射

一般说来, 不能定义相应于空间之间映射的只具上指标的张量的映射. 但如果给了映射 $F: x^i = x^i(x^{1'}, \cdots, x^{m'}), i = 1, \cdots, n$, 则可定义在点 $(x^{1'}, \cdots, x^{m'})$ 的向量空间到在点 $x^i(x^{1'}, \cdots, x^{m'}), i = 1, \cdots, n$ 的向量空间的映射 F_*. 这个映射的构造为

$$(F_*T)^i\big|_{x^k = x^k(x^{1'}, \cdots, x^{m'})} = T^{i'} \frac{\partial x^i}{\partial x^{i'}}\Big|_{(x^{1'}, \cdots, x^{m'})}. \tag{6}$$

相类似地可以对 $(k, 0)$ 型的张量构建映射 F_*. 因此切空间映射按相同于映射 F 作用的方式进行. 常常称切空间的映射 F_* 为映射 F 的微分.

如果 F 是空间中的光滑函数, 则 F_* 为在每点的切空间到实直线 $\mathbb{R}$ 的线性映射, 即向量的线性形式. 这便是在通常意义下的微分 $dF = \left(\dfrac{\partial F}{\partial x^i}\right)$.

在向量场上映射 F_* 不能 (整体地) 定义: 如果点 P_1 和 P_2 在映射 F 下映到同一个点 P, 则从点 P 将发出两个向量:$F_*T(P_1)$ 和 $F_*T(P_2), T = (T^i)$.

如果光滑映射 F 是相互一一的, 且映射 F^{-1} 光滑, 则可以定义 F_*, 且为切空间之间 (在相应点上) 的同构. 称这样的映射 F 为微分同胚.

习题 22.1. 证明对于微分形式 ω_1, ω_2 和光滑映射 F 成立公式

$$F^*(\omega_1 \wedge \omega_2) = F^*(\omega_1) \wedge F^*(\omega_2).$$

§23. 向量场

1. 微分同胚的单参数群

设在空间区域中给出了向量场, 它在坐标 $(x^1, \cdots, x^n)$ 下分量为 $\xi^i = \xi^i(x^1, \cdots, x^n)$. 每个这样的向量场都相伴了一个自治微分方程组

$$\dot{x}^i(t) = \xi^i(x^1(t), \cdots, x^n(t)), \quad i = 1, \cdots, n.$$
$$\dot{x}^i = \frac{dx^i}{dt}. \tag{1}$$

称它的解 $x^i = x^i(t)$ 为向量场 ξ^i 的积分曲线. 以

$$F_t^i(x_0^1, \cdots, x_0^n) = x^i = x^i(t, x_0^1, \cdots, x_0^n) \tag{2}$$

表示向量场 ξ^i 的以

$$x^i|_{t=0} = x_0^i \tag{3}$$

为初始条件的积分曲线.(由常微分方程的理论知, 如果函数 $\xi^i(x^1, \cdots, x^n)$ 光滑, 则存在满足 (3) 的唯一的解 (2).)

公式 (2) 给出了我们的区域到自己的映射

$$F_t: (x_0^1, \cdots, x_0^n) \longmapsto (x^1(t, x_0^1, \cdots, x_0^n), \cdots, x^n(t, x_0^1, \cdots, x_0^n)), \tag{4}$$

它依赖于参数 t (沿积分曲线随时间 t 移动). 由常微分方程理论知道, 映射 F_t 在已知点 $(x_0^1, \cdots, x_0^n)$ 的邻域中, 对小的 t 有定义, 并且局部为微分同胚. 另外,F_0 为恒同映射, 而微分同胚 F_t 构成了局部群:

$$F_{t+s} = F_t \circ F_s, \quad F_{-t} = (F_t)^{-1}. \tag{5}$$

"局部群"这个词表示, 如果这些等式两端对相应的参数值 $\pm t, s, t+s$ 有定义, 则这些等式成立. 我们得到了局部单参数的微分同胚群, 它与场 ξ^i 相关联.

当 t 小时, 映射 F_t 的显式表达为

$$x^i(t, x_0^1, \cdots, x_0^n) = x_0^i + t\xi^i(x_0^1, \cdots, x_0^n) + o(t). \tag{6}$$

在这个精确度下, 映射 F_t 的雅可比矩阵具形式

$$\frac{\partial x^i(t)}{\partial x_0^j} = \delta_j^i + t\frac{\partial \xi^i}{\partial x_0^j} + o(t). \tag{7}$$

它的逆映射的雅可比矩阵有形式

$$\frac{\partial x_0^i}{\partial x^j} = \delta_j^i - t\frac{\partial \xi^i}{\partial x^j} + o(t). \tag{8}$$

反之, 如果我们有了微分同胚 $F_t = (F_t^1, \cdots, F_t^n)$ 的单参数局部群, 于是照此唯一地恢复了一个向量场

$$\xi^i = \frac{d}{dt}F_t^i\bigg|_{t=0}, \quad i = 1, \cdots, n. \tag{9}$$

称这个向量场为*速度场*.

例 23.1. 考虑有坐标 (x, y) 的平面中的单参数群, 它是由绕坐标原点旋转角 t 组成. 于是映射 F_t 有形式

$$x = x_0 \cos\ t - y_0 \sin\ t, \quad y = x_0 \sin\ t + y_0 \cos\ t. \tag{10}$$

我们有

$$\frac{dx}{dt}\bigg|_{t=0} = -y_0, \quad \frac{dy}{dt}\bigg|_{t=0} = x_0.$$

因此, 在笛卡儿坐标 x, y 下速度场 $\xi^i, i = 1, 2$ 为

$$\xi(x, y) = (-y, x). \tag{11}$$

这个场的积分曲线为圆 $x^2+y^2=$ 常数 (图 30).

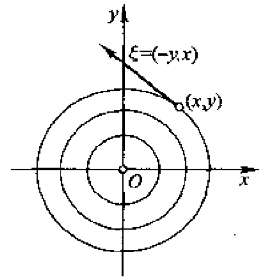


图 30

2. 向量场的指数映射

对应于向量场 $\xi(x)$ 的微分同胚 $F_t(x)$ 的单参数群作用于光滑函数 $f=f(x)$, 其规则为

$$(F_tf)(x)=f(F_t(x)). \tag{12}$$

例如, 我们考虑在直线上的单参数位移群 $F_t(x)=x+t$. 它的向量场 ξ 为常值向量. 变换 (12) 便为

$$(F_tf)(x)=f(x+t). \tag{13}$$

对于解析函数 $f(x)$ (即在任意点 x 的邻域中展开为收敛的幂级数), 对小的 t, 表达式 (13) 可以用泰勒级数表示

$$\begin{aligned}(F_tf)(x)&=f(x+t)=f(x)+tf'(x)+\frac{t^2}{2}f''(x)+\cdots\\&=\left(\sum_{n=0}^{\infty}\frac{t^n}{n!}\frac{d^n}{dx^n}\right)f(x)=\exp\left(t\frac{d}{dx}\right)f(x).\end{aligned} \tag{14}$$

推广这个计算结果, 我们有

定义 23.1. 称算子

$$\exp(t\partial_\xi)=1+t\partial_\xi+\frac{t^2}{2}(\partial_\xi)^2+\cdots=\sum_{n=0}^{\infty}\frac{t^n}{n!}(\partial_\xi)^n \tag{15}$$

为向量场 ξ 的指数算子,其中 ∂_ξ 按场 ξ 的方向的导数 (公式 (17.16)). 这个算子在函数 $f(x)$ 上的作用由等式

$$\exp(t\partial_\xi)f(x)=f(x)+t\partial_\xi f(x)+\frac{t^2}{2}(\partial_\xi)^2f(x)+\cdots. \tag{16}$$

定义. 这个表达式对那些使右端的级数收敛的函数 $f(x)$ 和 t 有定义.

命题 23.1. 对解析向量场 $\xi(x)$ (即所有的函数 $\xi^i(x^1,\cdots,x^n)$ 是解析的) 和解析函数 $f(x)$, 向量场 $\xi(x)$ 的指数算子对于充分小的 t 重合于算子 (12):

$$\exp(t\partial_\xi)f(x)=f(F_t(x)). \tag{17}$$

证明 考虑这样的点 $x=(x^1,\cdots,x^n)$, 使场 $\xi(x)$ 不化为零. 根据方程组 (1) 解的存在和唯一性定理, 在此点的某个邻域内可以作坐标变换

$$x^i=x^i(y^1,\cdots,y^n),\quad i=1,\cdots,n, \tag{18}$$

使得在新坐标下方程组 (1) 可以写为

$$\dot{y}^1 = 1, \quad \dot{y}^2 = 0, \cdots, \quad \dot{y}^n = 0 \tag{19}$$

(参看 [30]). 对于具解析右端的方程组, 变换 (18) 可以选成是解析的. 形如 $\widetilde{f}(y) = f(x(y))$ 的函数, 当 $f(x)$ 为解析时也是解析的. 在新坐标下, 等式 (17) 有形式

$$\exp\left(t\frac{\partial}{\partial y^1}\right)\widetilde{f}(y) = \widetilde{f}(y^1 + t, y^2, \cdots, y^n).$$

这可直接由 (14) 得到. 命题证完. □

习题 23.1. 计算算子 $\exp\left[(ax+b)\dfrac{d}{dx}\right]$.

3. 李导数. 例子

设 $\xi = (\xi^i)$ 为向量场,F_t 为对应的微分同胚的单参数群,$T^{i_1\cdots i_p}_{j_1\cdots j_q}$ 为 (p,q) 型张量. 由于映射的相互一一性,F_t 定义了张量由坐标 $x^i(t)$ 到 x^i_0 的变换规律. 对于张量 $T^{i_1\cdots i_p}_{j_1\cdots j_q}$ 从点 $(x^1(t), \cdots, x^n(t))$ 转移到点 $(x^1_0, \cdots, x^n_0)$, 我们得到了表达式

$$(F_tT)^{i_1\cdots i_p}_{j_1\cdots j_q} = T^{l_1\cdots l_p}_{k_1\cdots k_q}\frac{\partial x^{k_1}}{\partial x^{j_1}_0}\cdots\frac{\partial x^{k_q}}{\partial x^{j_q}_0}\frac{\partial x^{i_1}_0}{\partial x^{l_1}}\cdots\frac{\partial x^{i_p}_0}{\partial x^{l_p}}, \tag{20}$$

定义 23.2. 张量 $T^{i_1\cdots i_p}_{j_1\cdots j_q}$ 沿向量场 ξ 的李导数是指表达式

$$L_\xi T^{i_1\cdots i_p}_{j_1\cdots j_q} = \frac{d}{dt}(F_tT)^{i_1\cdots i_p}_{j_1\cdots j_q}\Big|_{t=0}. \tag{21}$$

于是, 李导数度量了张量 $T^{i_1\cdots i_p}_{j_1\cdots j_q}$ 变化的速度①, 这种变化是映射 F_t 给出的空间形变引起的. 因为 F_tT 对每个 t 值是 (p,q) 型的张量, 从而 $L_\xi T$ 也是 (p,q) 型的张量.

我们想要得到李导数的显式公式. 为此, 利用公式 (8) 和 (9) 重写 (20) 到 $o(t)$ 的精度. 我们有

$$\begin{aligned}
&(F_tT)^{i_1\cdots i_p}_{j_1\cdots j_q}\\
&= T^{l_1\cdots l_p}_{k_1\cdots k_q}\left(\delta^{k_1}_{j_1} + t\frac{\partial\xi^{k_1}}{\partial x^{j_1}_0}\right)\cdots\left(\delta^{k_q}_{j_q} + t\frac{\partial\xi^{k_q}}{\partial x^{j_q}_0}\right)\left(\delta^{i_1}_{l_1} - t\frac{\partial\xi^{i_1}}{\partial x^{l_1}}\right)\cdots\left(\delta^{i_p}_{l_p} - t\frac{\partial\xi^{i_p}}{\partial x^{l_p}}\right)\\
&= T^{i_1\cdots i_p}_{j_1\cdots j_q} + t\left[T^{i_1\cdots i_p}_{kj_2\cdots j_q}\frac{\partial\xi^k}{\partial x^{j_1}_0} + \cdots + T^{i_1\cdots i_p}_{j_1\cdots j_{q-1}k}\frac{\partial\xi^k}{\partial x^{j_q}_0}\right.\\
&\qquad\left.-T^{li_2\cdots i_p}_{j_1\cdots j_q}\frac{\partial\xi^{i_1}}{\partial x^l} - \cdots - T^{i_1\cdots i_{p-1}l}_{j_1\cdots j_q}\frac{\partial\xi^{i_p}}{\partial x^l}\right] + o(t).
\end{aligned}\tag{22}$$

①在连续介质力学的理论中, 称表达式 $dT/d\tau = \partial T/\partial\tau + L_\xi T$ 为 T 沿速度场 ξ 的 "全导数", 这里的 $T = T(\tau, x)$ 为任意向量场.

对 (22) 作关于 t 的微分并令 $t=0$, 得到

$$L_\xi T^{i_1\cdots i_p}_{j_1\cdots j_q}=\xi^s\frac{\partial T^{i_1\cdots i_p}_{j_1\cdots j_q}}{\partial x^s}+T^{i_1\cdots i_p}_{kj_2\cdots j_q}\frac{\partial\xi^k}{\partial x^{j_1}}+\cdots+T^{i_1\cdots i_p}_{j_1\cdots j_{q-1}k}\frac{\partial\xi^k}{\partial x^{j_q}}-$$

$$T^{li_2\cdots i_p}_{j_1\cdots j_q}\frac{\partial\xi^{i_1}}{\partial x^l}-\cdots-T^{i_1\cdots i_{p-1}l}_{j_1\cdots j_q}\frac{\partial\xi^{i_p}}{\partial x^l}. \tag{23}$$

例 23.2. 零阶张量即标量. 我们有

$$L_\xi f=\xi^i\frac{\partial f}{\partial x^i}=\partial_\xi f. \tag{24}$$

我们得到函数 f 对方向 ξ 的方向导数 (即函数沿向量场的导数). 如果 $L_\xi f=0$, 则函数 f 沿向量场 ξ 的积分曲线为常值. 这样的函数被称为场 ξ 的积分或者称为对应的方程组 $\dot{x}^i=\xi^i(x^1,\cdots,x^n)$ 的积分, 例如在第 1 小节中考虑过的例子里, 函数 $f(x,y)=x^2+y^2$ 便是积分.

如果 f 为向量场 ξ 的积分, 则这个场的积分曲线完全位于水平面 $f(x^1,\cdots,x^n)=$ 常数中; 显然, 这个场 ξ 切于这些曲面. 这允许把原来的方程组 $\dot{x}^i=\xi^i(x^1,\cdots,x^n)$ 的阶从 n 减少到 $n-1$, 这是由将它限制 (即限制向量场 $\xi^i(x^1,\cdots,x^n)$) 在超曲面 $f(x^1,\cdots,x^n)=$ 常数上. 我们知道, 这个超曲面的维数等于 $n-1$. 就是说, 向量场 ξ 的限制这个简单的几何过程对应了在微分方程理论中所熟知的命题, 即给出了方程组的首次积分则可降低其阶数 1.

例 23.3. (1,0) 型张量, 即向量. 设 $\eta=(\eta^i)$ 为向量场. 因此由公式 (23) 我们得到

$$L_\xi\eta^i=\xi^j\frac{\partial\eta^i}{\partial x^j}-\eta^j\frac{\partial\xi^i}{\partial x^j}. \tag{25}$$

由此得到

$$L_\xi\eta=-L_\eta\xi. \tag{26}$$

我们要解释向量场 $L_\xi\eta$ 是如何作用于函数的. 我们有重要的定理.

定理 23.1. $\partial_{L_\xi\eta}f=\partial_\xi(\partial_\eta f)-\partial_\eta(\partial_\xi f)=[\partial_\xi,\partial_\eta]f$. 所以, 算子 ∂_ξ 和 ∂_η 的换位子仍是一阶微分算子, 即向量场 $L_\xi\eta=-L_\eta\xi$.

证明 显式地计算换位子 $[\partial_\xi,\partial_\eta]$

$$\begin{aligned}[\partial_\xi,\partial_\eta]f&=\partial_\xi(\partial_\eta f)-\partial_\eta(\partial_\xi f)=\partial_\xi\left(\eta^i\frac{\partial f}{\partial x^i}\right)-\partial_\eta\left(\xi^i\frac{\partial f}{\partial x^i}\right)\\&=\xi^j\eta^i\frac{\partial^2 f}{\partial x^i\partial x^j}+\xi^j\frac{\partial\eta^i}{\partial x^j}\frac{\partial f}{\partial x^i}-\eta^j\xi^i\frac{\partial^2 f}{\partial x^i\partial x^j}-\eta^j\frac{\partial\xi^i}{\partial x^j}\frac{\partial f}{\partial x^i}\\&=\left(\xi^j\frac{\partial\eta^i}{\partial x^j}-\eta^j\frac{\partial\xi^i}{\partial x^j}\right)\frac{\partial f}{\partial x^i}=\partial_{L_\xi\eta}f;\end{aligned}$$

这里 $\xi^j\eta^i\frac{\partial^2 f}{\partial x^i\partial x^j}-\xi^i\eta^j\frac{\partial^2 f}{\partial x^i\partial x^j}=0$, 是因为 $\frac{\partial^2 f}{\partial x^i\partial x^j}=\frac{\partial^2 f}{\partial x^j\partial x^i}$. 证完. □

定义 23.3. 称向量场 $[\xi,\eta]=L_\xi\eta$ 为向量场 ξ 和 η 的换位子.

我们注意有莱布尼茨公式的完全类比的公式

$$L_\xi(f\eta)=fL_\xi\eta+\eta(\partial_\xi f).$$

利用公式 (25), 由直接计算便得出它的证明.

由此公式立即推导出下面的定理.

定理 23.2. 设在 $\mathbb{R}^n$ 中给出了一组光滑向量场 $\xi_1,\cdots,\xi_k$. 为了存在一个坐标系 $y^1,\cdots,y^n$, 使得向量 ξ_j 切于坐标轴 $(y^j),j=1,\cdots,k$, 必须满足下面条件

$$[\xi_j,\xi_k]=f_{jk}^{(1)}\xi_j+f_{jk}^{(2)}\xi_k, \tag{27}$$

其中 $f_{jk}^{(i)}$ 为点的标量函数.

证明 如果向量场 ξ_j 为坐标系的基坐标向量 e_j, 于是, 因为 $\partial_{e_i}=\dfrac{\partial}{\partial y^i}$, 且 $\dfrac{\partial^2}{\partial y^j\partial y^k}=\dfrac{\partial^2}{\partial y^k\partial y^j}$, 故 $[\xi_j,\xi_k]=0$. 如果 $\xi_j=f_j(y)e_j$, e_j 为基坐标向量, 而 $f_j(y)$ 为点的光滑函数, 于是

$$[\xi_j,\xi_k]=[f_je_j,f_ke_k]=f_j\frac{\partial f_k}{\partial y^j}e_k-f_k\frac{\partial f_j}{\partial y^k}e_j.$$

令 $f_{jk}^{(1)}=-\dfrac{f_k}{f_j}\dfrac{\partial f_j}{\partial y^k}, f_{jk}^{(2)}=\dfrac{f_j}{f_k}\dfrac{\partial f_k}{\partial y^j}$, 便得到所要求的关系. □

例 23.4. (0,1) 型张量 T_j, 即余向量. 根据公式 (23) 有

$$(L_\xi T)_j=\xi^k\frac{\partial T_j}{\partial x^k}+T_k\frac{\partial\xi^k}{\partial x^j}. \tag{28}$$

对于函数的微分 $T_j=\dfrac{\partial f}{\partial x^j}$, 我们得到

$$(L_\xi T)_j=\xi^k\frac{\partial^2 f}{\partial x^k\partial x^j}+\frac{\partial f}{\partial x^k}\frac{\partial\xi^k}{\partial x^j}. \tag{29}$$

结论 取李导数 L_ξ 和取微分可交换

$$L_\xi(df)=d(L_\xi f). \tag{30}$$

例 23.5. (0,2) 型张量 g_{ij}, 即双线性形式. 我们有

$$L_\xi g_{ij}=\xi^s\frac{\partial g_{ij}}{\partial x^s}+g_{kj}\frac{\partial\xi^k}{\partial x^i}+g_{ik}\frac{\partial\xi^k}{\partial x^j}\equiv u_{ij}. \tag{31}$$

称张量 u_{ij} 为 (小) 形变张量. 它描述了空间的度量 g_{ij} 在小形变 F_t 下的变化, 其中的 F_t 由向量场 ξ 定义. 特别, 如果度量是欧氏的,$g_{ij}=\delta_{ij}$, 则张量 u_{ij} 有形式 (见 §17)

$$u_{ij}=\frac{\partial\xi^i}{\partial x^j}+\frac{\partial\xi^j}{\partial x^i}. \tag{32}$$

例 23.6. 计算体积元 $\sqrt{|g|}\varepsilon_{i_1\cdots i_n}$(或 $\sqrt{|g|}dx^1\wedge\cdots\wedge dx^n$) 的李导数. 根据公式 (23),

$$L_\xi\sqrt{|g|}\varepsilon_{i_1\cdots i_n}=\xi^k\frac{\partial\sqrt{|g|}}{\partial x^k}\varepsilon_{i_1\cdots i_n}+\sqrt{|g|}\left(\varepsilon_{ki_2\cdots i_n}\frac{\partial\xi^k}{\partial x^{i_1}}+\cdots+\varepsilon_{i_1\cdots i_{n-1}k}\frac{\partial\xi^k}{\partial x^{i_n}}\right). \quad (33)$$

公式 (33) 的右端括号里面的表达式等于 $\frac{\partial\xi^i}{\partial x^i}\varepsilon_{i_1\cdots i_n}$(见习题 18.5). 故而 (33) 的右端等于

$$\begin{aligned}L_\xi\sqrt{|g|}\varepsilon_{i_1\cdots i_n}&=\sqrt{|g|}\varepsilon_{i_1\cdots i_n}\left(\xi^k\frac{\partial\ln\sqrt{|g|}}{\partial x^k}+\frac{\partial\xi^i}{\partial x^i}\right)\\&=\sqrt{|g|}\varepsilon_{i_1\cdots i_n}\left(\xi^k\frac{1}{2}g^{im}\frac{\partial g_{im}}{\partial x^k}+\frac{\partial\xi^i}{\partial x^i}\right).\end{aligned}$$

然而右端括号中为形变张量的迹 $\mathrm{Tr}(u_{im})=g^{im}u_{im}$, 这里的 u_{im} 由公式 (31) 定义. 于是最后得到了体积元的李导数的表达式

$$L_\xi(\sqrt{|g|}\varepsilon_{i_1\cdots i_n})=\frac{1}{2}g^{im}u_{im}\sqrt{|g|}\varepsilon_{i_1\cdots i_n}, \quad (34)$$

其中右端为形变张量 u_{im} 的迹. 在欧氏情形 $g_{ij}=\delta_{ij}$ 时, 形变张量的迹为

$$\frac{1}{2}g^{im}u_{im}=\frac{\partial\xi^i}{\partial x^i}.$$

习题

23.2. 证明李导数的莱布尼茨公式

$$L_\xi(T\otimes R)=(L_\xi T)\otimes R+T\otimes(L_\xi R),$$

其中 T,R 为任意两个张量.

23.3. 设 ω_1,ω_2 为两个微分形式. 证明

$$L_\xi(\omega_1\wedge\omega_2)=(L_\xi\omega_1)\wedge\omega_2+\omega_1\wedge(L_\xi\omega_2).$$

23.4. 设 F 为区域 U 到区域 V 内的微分同胚,X_1,X_2 为 U 上向量场,$Y_i=F_*(X_i)$ 为 V 上的相对应的向量场. 证明 $F_*[X_1,X_2]=[Y_1,Y_2]$.

23.5. 设 F_t,G_s 为微分同胚的单参数群,X,Y 为其相应的向量场. 证明微分同胚 F_t 和 G_s 对任意的 t 和 s 可交换当且仅当 X 和 Y 的换位子等于零.

23.6. 设 $X_1,\cdots,X_n$ 是在 n 维区域中线性无关的向量场, 其中 $[X_i,X_j]=0$. 证明, 存在 (局部的) 坐标系 $(x^1,\cdots,x^n)$ 使向量场 X_i 切于第 i 个坐标轴:$\partial_{X_i}(x^k)=\delta_i^k$.

§24. 李代数

1. 李代数和向量场

定义 24.1. 在向量空间 V 中给出了一个反称的双线性算子 $[,]$, 如果它满足雅可比恒等式

$$[\xi,[\eta,\zeta]]+[\eta,[\zeta,\xi]]+[\zeta,[\xi,\eta]]=0, \tag{1}$$

则称其为李代数.

注 对任意 $\xi\in V$ 我们引进算子 $\mathrm{ad}\,\xi$, 这是个线性映射 $\mathrm{ad}\,\xi:V\to V$; 令 $\mathrm{ad}\,\xi(\eta)=[\xi,\eta]$. 雅可比恒等式表明映射 $\mathrm{ad}\,\xi$ 就像在代数中所说的那样, 是李代数 V 的"导子"(即满足莱布尼茨公式)

$$\mathrm{ad}\,\xi([\eta,\zeta])=[\mathrm{ad}\,\xi(\eta),\zeta]+[\eta,\mathrm{ad}\,\xi(\zeta)]. \tag{2}$$

例

24.1. 三维欧氏空间在向量积的运算下是个李代数.

24.2. 设 V 为某个线性算子的代数, 那么 V 可以转化为李代数, 只要令

$$[A,B]=AB-BA. \tag{3}$$

可以证明对这样的括号雅可比恒等式成立. 我们有

$$\begin{aligned}
&[A,[B,C]]=A[B,C]-[B,C]A=ABC-ACB-BCA+CBA,\\
&[C,[A,B]]=CAB-CBA-ABC+BAC,\\
&[B,[C,A]]=BCA-BAC-CAB+ACB.
\end{aligned} \tag{4}$$

把这些表达式相加等于零.

注 在李代数中算子 $[,]$ 常被称做换位子.

推论 1 所有 n 阶矩阵的空间 $M(n,\mathbb{R})$ 关于换位子

$$[A,B]=AB-BA \tag{5}$$

是一个李代数.

推论 2 在空间 $\mathbb{R}^n$ 的区域中的向量场关于换位子

$$[\xi,\eta]^i=\xi^j\frac{\partial\eta^i}{\partial x^j}-\eta^j\frac{\partial\xi^i}{\partial x^j} \tag{6}$$

构成一个李代数.

推论 2 由定理 23.1 推出. 向量场的李代数被理解为是无穷维的.

设在空间区域中给出了某个曲面及切于这个曲面的向量场 ξ,η. 于是成立

定理 24.1. 如果两个向量场 ξ,η 切于某个光滑的曲面, 则它们的换位子也切于这个曲面.

证明 因为任意曲面可以 (局部地) 表示为超曲面的交集, 故而只要对超曲面 $f(x^1,\cdots,x^n)=0$ 证明定理就足够了. 不失一般性, 可以认为此曲面由方程

$$x^n = 0 \tag{7}$$

给出. 场 ξ,η 切于此曲面表明

$$\partial_\xi f|_{f=0} = \partial_\eta f|_{f=0} = 0.$$

对于形如 (7) 的曲面, 这个相切的条件为

$$\xi^n|_{x^n=0} = 0, \quad \eta^n|_{x^n=0} = 0.$$

于是对于换位子 $[\xi,\eta]$ 我们有

$$[\xi,\eta]^n = \xi^k \frac{\partial \eta^n}{\partial x^k} - \eta^k \frac{\partial \xi^n}{\partial x^k}. \tag{8}$$

然而如果 $x^n=0$, 则当 $k\neq n$ 时有 $\dfrac{\partial \xi^n}{\partial x^k} = \dfrac{\partial \eta^n}{\partial x^k} = 0$. 故而 $[\xi,\eta]^n|_{x^n=0}=0$. 定理得证. □

推论 切于某个光滑曲面的向量场构成了所有向量场的李代数的子代数.

2. 基本的矩阵李代数

前面考察过的每个线性变换群都关联了矩阵李代数. 群的单位元处的切空间就是这个李代数的空间; 换位子即通常矩阵的换位子.

我们来列出最重要的矩阵群和它们在单位元的切空间. 在单位元的切空间将用与群的记号相同的记号表示, 但用小写的字母.

1) 特殊线性群 $SL(n,\mathbb{R})$ (或 $SL(n,\mathbb{C})$) 为行列式为 1 的 n 阶实 (复) 矩阵群. 在单位的切空间 $sl(n,\mathbb{R})$ (或 $sl(n,\mathbb{C})$) 为其迹等于零的矩阵空间.

2) 旋转群$SO(n,\mathbb{R})$ (或 $SO(n,\mathbb{C})$) 为行列式为 1 的实 (或复) 的正交矩阵的群:

$$A^{\mathrm{T}}A = 1, \quad \det A = 1, \quad A\in SO(n,\mathbb{R})\ (SO(n,\mathbb{C})). \tag{9}$$

那么 $so(n,\mathbb{R}), so(n,\mathbb{C})$ 为反称矩阵的代数 (实或复的)

$$X^{\mathrm{T}} = -X, \quad X\in so(n,\mathbb{R})\ \ (so(n,\mathbb{C})). \tag{10}$$

3) 伪正交群$SO(p,q)$. 设 $G=(g_{ij})$ 为空间 $\mathbb{R}^n_{p,q}, p+q=n$ 中的伪欧氏度量. 群 $SO(p,q)$ 为行列式为 1 的实矩阵 A 的群, 它保持形式 $G=(g_{ij})$ 不变

$$A^{\mathrm{T}}GA = G, \quad \det A = 1, \quad A\in SO(p,q). \tag{11}$$

这时 $so(p,q)$ 是那样一些矩阵 $X=(x^i_j)$ 构成的代数, 它们满足

$$GX + X^{\mathrm{T}}G = 0, \quad X\in so(p,q), \tag{12}$$

或者

$$g_{ij}x^j_k + x^j_i g_{jk} = 0. \tag{13}$$

最后面这个等式意味着矩阵 $GX = (u_{ik})$

$$u_{ik} = g_{ij}x^j_k \tag{14}$$

为反对称的. 这些等式建立了空间 $so(p,q)$ 与所有反称矩阵 ($n = p+q$ 阶) 的空间之间的同构.

4) 酉群$U(n)$, 即 n 阶酉矩阵的群

$$A^{\mathrm{T}}A = 1, \quad A \in U(n). \tag{15}$$

代数 $u(n)$ 由反埃尔米特矩阵组成

$$\bar{X}^{\mathrm{T}} = -X, \quad X \in u(n). \tag{16}$$

5) 特殊酉群$SU(n)$, 即行列式为 1 的酉矩阵群. 代数 $su(n)$ 由迹为 0 的反埃尔米特矩阵组成

$$\bar{X}^{\mathrm{T}} = -X, \quad \mathrm{Tr}\, X = 0, \quad X \in su(n). \tag{17}$$

6) 伪酉群$U(p,q)$, 即复 n 维空间中线性变换, 其中 $n = p+q$, 且它保持内积

$$\langle \xi, \eta \rangle = \sum_{i=1}^{p} x^i \bar{y}^i - \sum_{i=p+1}^{n} x^i \bar{y}^i = g_{ij}x^i \bar{y}^j; \tag{18}$$

$$\xi = (x^1, \cdots, x^n), \quad \eta = (y^1, \cdots, y^n); \quad n = p+q$$

不变.

如果 $G = (g_{ij})$ 为公式 (18) 中的矩阵, 于是对群 $U(p,q)$ 中的矩阵 A 有

$$A^{\mathrm{T}}GA = G, \quad A \in U(p,q). \tag{19}$$

从而空间 $u(p,q)$ 由那些矩阵 $X = (x^j_i)$ 组成, 它们满足

$$GX + \bar{X}^{\mathrm{T}}G = 0. \tag{20}$$

空间 $u(p,q)$ 同构地与反埃尔米特矩阵的空间相联系: 矩阵 $(x^j_k) \in u(p,q)$ 对应于反埃尔米特矩阵 (u_{ik})

$$u_{ik} = g_{ij}x^j_k, \quad \bar{u}_{ki} = -u_{ik}. \tag{21}$$

7) 群 $SU(p,q)$ 是群 $U(p,q)$ 的子群, 它由行列式为 1 的那些矩阵组成;$su(p,q)$ 中的矩阵具有零迹

$$x^i_i = 0 \Leftrightarrow g^{ik}u_{ik} = 0. \tag{22}$$

定理 24.2. 空间 $sl(n,\mathbb{R}), sl(n,\mathbb{C}), so(n,\mathbb{R}), so(n,\mathbb{C}), so(p,q), u(n), su(n), u(p,q), su(p,q)$ 在矩阵的换位子下是李代数.

证明 需要验证上面所列出的空间中每一个都在换位子下封闭. 我们证明

a) 如果 $\mathrm{Tr}\, X=0, \mathrm{Tr}\, Y=0$, 则 $\mathrm{Tr}[X,Y]=0$;

b) 如果 X,Y 满足条件 (12), 则 $[X,Y]$ 也满足这个条件;

c) 如果 X,Y 满足条件 (20), 从而 $[X,Y]$ 也满足.

我们有 $\mathrm{Tr}[X,Y]=\mathrm{Tr}(XY)-\mathrm{Tr}(YX)=0$ (对任意 X,Y). 论断 a) 得证.

设矩阵 X,Y 满足条件 (12), 即

$$X^{\mathrm{T}}G=-GX,\quad Y^{\mathrm{T}}G=-GY.$$

于是

$$[X,Y]^{\mathrm{T}}G=Y^{\mathrm{T}}X^{\mathrm{T}}G-X^{\mathrm{T}}Y^{\mathrm{T}}G=GYX-GXY=-G[X,Y].$$

故论断 b) 得证.

类似地证明论断 c). 定理证完. □

定义 24.2. 设 G 为 1)—7) 的变换群中的一个. 在群 G 的单位元处的切空间及在其上赋予了矩阵的换位子运算, 则称其为群 G 的李代数.

例 24.3. 三维空间旋转群的李代数 $so(3,\mathbb{R})$ 由三阶反称矩阵组成. 在这个矩阵空间中引进一组基 X_1,X_2,X_3, 它们是

$$X_1=\begin{pmatrix}0&0&0\\0&0&-1\\0&1&0\end{pmatrix},\quad X_2=\begin{pmatrix}0&0&1\\0&0&0\\-1&0&0\end{pmatrix},\quad X_3=\begin{pmatrix}0&-1&0\\1&0&0\\0&0&0\end{pmatrix}. \tag{23}$$

于是我们有

$$[X_1,X_2]=X_3,\quad [X_2,X_3]=X_1,\quad [X_3,X_1]=X_2. \tag{24}$$

结论 群 $SO(3,\mathbb{R})$ 的李代数同构于三维欧氏空间中的向量在向量积下的李代数.

例 24.4. 考虑李代数 $so(p,q)$. 设伪欧几里得度量的形式为

$$g_{ij}=\varepsilon_i\delta_{ij},\quad \varepsilon_i=\pm 1. \tag{25}$$

代数 $so(p,q)$ 作为反称矩阵而实现 (参看公式 (14)), 另外换位子有形式

$$\begin{aligned}&[u,v]_{ij}=\sum_k\varepsilon_k(u_{ik}v_{kj}-v_{ik}u_{kj}),\\&u=(u_{ij}),\quad v=(v_{ij}),\quad u_{ji}=-u_{ij},\quad v_{ji}=-v_{ij}.\end{aligned} \tag{26}$$

例 24.5. 李代数 $su(2)$. 在迹为 0 的反称 2×2 矩阵的空间中, 选取基 s_1, s_2, s_3, 其中

$$s_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad s_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad s_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \tag{27}$$

于是有

$$[s_1, s_2] = 2s_3, \quad [s_2, s_3] = 2s_1, \quad [s_3, s_1] = 2s_2. \tag{28}$$

由引理 14.6 知, 这个李代数同构于纯虚四元数 (即 $\bar{x} = -x$) 的李代数, 其中的换位子为 $[x, y] = xy - yx$, 在此同构下为

$$i \leftrightarrow s_1, \quad j \leftrightarrow s_2, \quad k \leftrightarrow s_3. \tag{29}$$

定理 24.3. 存在李代数间的同构

$$su(2) \simeq so(3, \mathbb{R}). \tag{30}$$

证明 对应于矩阵 $X \in su(2)$ 我们给出三维空间 $su(2)$ 的线性变换 ad X :

$$Z \mapsto \text{ad } X(Z) = [X, Z], \quad X, Z \in su(2). \tag{31}$$

这时由雅可比恒等式有

$$[\text{ad } X, \text{ad } Y] = \text{ad}[X, Y]. \tag{32}$$

这表明, 映射 (31) 为李代数 $su(2)$ 到空间 $su(2)$ 的线性算子的李代数之间的同态. 空间 $su(2)$ 是欧几里得的; 向量 $Z = bs_1 + cs_2 + ds_3 \leftrightarrow bi + cj + dk$ 的长度等于

$$|Z|^2 = b^2 + c^2 + d^2 = \det Z. \tag{33}$$

变换

$$Z \mapsto AZA^{-1}, \quad A \in SU(2) \tag{34}$$

为正交, 其意义是在内积 (33) 下的 (参看 §14)

$$\det(AZA^{-1}) = \det Z.$$

引理 24.1. 变换 ad $X, X \in su(2)$, 在度量 (33) 下为反称.

证明 设群 $SU(2)$ 中变换族 $A = A(t)$, 使得

$$\left.\frac{dA(t)}{dt}\right|_{t=0} = X, \quad A(0) = 1. \tag{35}$$

然后将此变换族

$$Z \mapsto A(t)ZA(t)^{-1}, \quad Z \in su(2)$$

对 t 求导, 当 $t = 0$ 有

$$Z \mapsto XZ - ZX = \text{ad } X(Z).$$

因此, 这个变换是反称的 (参看 §5.3). 引理得证. □

因为我们有同态

$$su(2) \to so(3,\mathbb{R}), \quad X \mapsto \text{ad } X. \tag{36}$$

它的核等于零 (如果 ad $X(Z) = 0$ 对所有 Z 成立, 则 $X = 0$), 但空间 $su(2)$ 和 $so(3,\mathbb{R})$ 的维数相等 (等于 3). 定理得证. □

注 在基 s_1, s_2, s_3 (27) 下变换 ad s_1, ad s_2, ad s_3 的矩阵为

$$\text{ad } s_1 = 2X_1, \quad \text{ad } s_2 = 2X_2, \quad \text{ad } s_3 = 2X_3, \tag{37}$$

其中基 X_1, X_2, X_3 在空间 $so(3,\mathbb{R})$ 由公式 (23) 给出.

例 24.6. 李代数 $sl(2,\mathbb{R})$. 在此代数中引进矩阵基 Y_0, Y_1, Y_2 为

$$Y_0 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad Y_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad Y_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{38}$$

这些矩阵的换位子为

$$[Y_0, Y_1] = -2Y_2, \quad [Y_0, Y_2] = 2Y_1; \quad [Y_1, Y_2] = 2Y_0. \tag{39}$$

定理 24.4. 存在李代数的同构

$$sl(2,\mathbb{R}) \simeq so(1,2).$$

证明 像在定理 24.3 那样,$sl(2,\mathbb{R})$ 中每个矩阵 Y 对应于三维空间 $sl(2,\mathbb{R})$ 的线性变换 ad Y. 空间 $sl(2,\mathbb{R})$ 到自身的变换

$$Y \mapsto AYA^{-1} \tag{40}$$

保持二次型

$$|Y|^2 = \det Y \tag{41}$$

不变. 因此 (参看引理 1) 变换 ad Y 在度量 (41) 的意义下为反称. 但是这个伪欧氏度量为 (1,2) 型

$$\begin{aligned}\det Y &= \det(y^0 Y_0 + y^1 Y_1 + y^2 Y_2) \\ &= \det \begin{pmatrix} y_1 & y_2 + y_0 \\ y_2 - y_0 & -y_1 \end{pmatrix} = y_0^2 - y_1^2 - y_2^2. \end{aligned} \tag{42}$$

于是我们把 ad Y 实现为 $so(1,2)$ 中元. 定理证完. □

3. 线性向量场

设 $X = (X^i_k)$ 为 n 阶实 (或复) 矩阵. 我们在空间 $\mathbb{R}^n$ (或 $\mathbb{C}^n$) 中构造一个向量场 T_X : 令在点 $x \in \mathbb{R}^n$(或 $x \in \mathbb{C}^n$) 的值等于

$$T_X(x) = -Xx, \tag{43}$$

或者在坐标下为

$$T^i_X(x) = -X^i_k x^k. \tag{44}$$

定义 24.3. 称 (43) 型的向量场为线性向量场.

我们来求线性向量场 (43) 的积分曲线. 由定义, 我们有常系数的线性微分方程组 (向量形式)

$$\dot{x} = -Xx. \tag{45}$$

定理 24.5. 向量场 (43) 的积分曲线 $x(t)$, 在满足初始条件 $x(0) = x_0$ 时为

$$x(t) = \exp(-tX)x_0. \tag{46}$$

证明 我们要验证曲线 (46) 满足微分方程 (45). 想到 (参看 §14.2) 矩阵的指数算子的定义是和

$$\exp(-tX) = 1 - \frac{tX}{1!} + \frac{t^2X^2}{2!} + \cdots. \tag{47}$$

于是将此级数对 t 微分便得到

$$\frac{d}{dt}\exp(-tX) = -X + \frac{tX^2}{1!} - \cdots = -X\exp(-tX). \tag{48}$$

故

$$\frac{dx}{dt} = \frac{d}{dt}(\exp(-tX)x_0) = -X\exp(-tX)x_0 = -Xx,$$

即满足关系 (45). 于是由微分方程解的唯一性便得到了我们的定理. □

由此定理知道, 由线性场 T_X 生成的微分同胚单参数群是乘以矩阵 $\exp(-tX)$.

例 24.7. 矩阵 X_1, X_2, X_3 构成了李代数 $so(3)$ 的基 (23), 它们生成了三维欧氏空间的三个线性向量场. 我们以 L_x, L_y, L_z 表示它们. 这些向量场在坐标为 (x, y, z) 点上的值等于

$$L_x = (0, +z, -y), \quad L_y = (-z, 0, +x), \quad L_z = (+y, -x, 0). \tag{49}$$

这些向量场对应了三个单参数群: 分别为空间 $\mathbb{R}^3$ 的绕 x, y, z 轴的旋转.

设 X 和 Y 为两个 n 阶矩阵. 我们来计算两个线性场 T_X 和 T_Y 的换位子.

定理 24.6. 向量场 T_X 和 T_Y 的换位子为

$$[T_X, T_Y] = T_{[X,Y]}, \tag{50}$$

其中 $[X, Y] = XY - YX$ 为矩阵 X, Y 的换位子.

证明 利用在坐标表示下向量场的换位子的公式 (8), 我们有

$$\begin{aligned}[T_X, T_Y]^i &= X_l^k x^l \frac{\partial(Y_m^i x^m)}{\partial x^k} - Y_l^k x^l \frac{\partial(X_m^i x^m)}{\partial x^k} \\ &= X_l^k x^l Y_k^i - Y_l^k x^l X_k^i = (-[X, Y]x)^i.\end{aligned}$$

然而最后一个表达式为线性向量场 $T_{[X,Y]}$ 的第 i 个分量. 定理证完. □

推论 线性向量场在通常的换位子下组成了有限维李代数, 它同构于所有 n 阶矩阵的李代数.

设 G 为我们在第 2 小节中考虑过的 n 维空间 (实和复) 的变换群中的一个, $\mathfrak{g}$ 为它的李代数. 形如 T_X (其中 X 为 $\mathfrak{g}$ 中矩阵) 的线性向量场构成了一个李代数, 它同构于这个群的李群. 对应的微分同胚的单参数李群由乘以群 G 的单参数子群的元素得到.

注 对应于线性向量场 $T_X, X \in \mathfrak{g}$, 我们有微分算子 ∂_{T_X}, 它定义在 $\mathbb{R}^n$ 的光滑函数上; 我们称其为群 G 的*作用生成元*. 知道了这些生成元, 就可以重新建立群 G 在函数上的作用: 取向量场 $\xi = T_X$ 的指数算子 (见 §23)

$$f(F_t(x)) = f(\exp(-tX)x) = \exp(-t\partial_\xi)f(x). \tag{51}$$

例 24.8. 三维空间的旋转群 $SO(3)$ 的生成元为微分算子

$$L_x = z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}, \quad L_y = x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x}, \quad L_z = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}. \tag{52}$$

根据定理 24.6 这些微分算子的换位子有下面公式:

$$[L_x, L_y] = L_z, \quad [L_y, L_z] = L_x, \quad [L_x, L_z] = L_y. \tag{53}$$

4. 变换群上的左不变向量场

设 X 为一个固定的 n 阶矩阵 (按定义为实矩阵). 它对应于矩阵空间 $\mathbb{R}^{n^2}$ 的线性变换

$$A \mapsto AX. \tag{54}$$

我们以 L_X 表示在 n 阶矩阵的空间 $\mathbb{R}^{n^2}$ 中的线性向量场. 向量场 L_X 在点 A(即 $n \times n$ 矩阵) 上的取值等于

$$L_X(A) = AX. \tag{55}$$

向量场 L_X 的积分曲线可按定理 24.5 求出. 按定义我们有微分方程组 (是一个矩阵方程)

$$\dot{A} = L_X(A) = AX. \tag{56}$$

这个方程组在初始条件 $A|_{t=0} = A_0$ 下的解为

$$A = A_0 \exp(tX). \tag{57}$$

因此, 由向量场 L_X 生成的微分同胚的单参数群等于以矩阵 $\exp(tX)$ 作右乘积 (以 $\exp\ tX$ 作右位移).

按定理 24.6, 形如 L_X 的向量场的换位子的表达式为

$$[L_X, L_Y] = L_{[X,Y]}. \tag{58}$$

向量场 L_X 具有重要的性质: 左不变性 (关于左位移的不变性):

$$BL_X(A) = L_X(BA). \tag{59}$$

设 G 为在第 2 小节中所考虑过的变换群之一, 它可作为所有 n 阶矩阵的空间 $\mathbb{R}^{n^2}$ 中的一个光滑曲面. 设 $\mathfrak{g}$ 为群 G 在单位元的切空间. 我们已知 (§14.2), 如果 $X \in \mathfrak{g}$, 则矩阵 $\exp(tX)$ 对任意的 t 都在 G 中 (G 中的单参数子群).

引理 24.2. 如果 $X \in \mathfrak{g}$, 则向量场 L_X 切于曲面 G, 即给出了群 G 上的向量场.

证明 在点 $A \in G$ 处向量 $L_X(A)$ 为曲线 $A\exp\ tX \in G$ 的初始速度向量. 引理得证. □

我们把向量场 L_X 在曲面 G 上的限制仍然记作 L_X, 这里的 X 为群 G 的李代数 $\mathfrak{g}$ 中的矩阵. 注意, 向量场 L_X 在群的单位元的值等于 X; 场 L_X 在 G 上具有对于左位移 G 中元素不变的性质.

定义 24.4. 称群 G 上形如 L_X 的向量场为群 G 上的*左不变向量场*,其中 $X \in \mathfrak{g}$ 为这个群的李代数中元.

由定理 24.1 和公式 (58) 立即得到

定理 24.7. 群 G 上的左不变向量场构成了李代数, 它同构于群 G 的李代数 $\mathfrak{g}$.

下面的引理对我们也是有用的:

引理 24.3. 左不变向量场在群 G 的每点的值构成了群 G 在这个点的整个切空间.

证明 如果 $X_1, \cdots, X_N$ 为李代数 $\mathfrak{g}$ 的基, 则向量 $L_{X_1}(A), \cdots, L_{X_N}(A)$ 在每点 $A \in G$ 线性无关, 从而组成了切空间的基. 证完. □

5. 基灵度量

我们首先在李代数 $\mathfrak{g}$ 上引进基灵度量的概念.

定义 24.5. 称李代数 $\mathfrak{g}$ 上欧氏或伪欧氏内积 $\langle , \rangle_0$ 为基灵度量是说, 所有形如 $\mathrm{ad}\ X$ 的算子在这个度量下是反称的:

$$\langle \mathrm{ad}\ X(Y), Z\rangle_0 = -\langle Y, \mathrm{ad}\ X(Z)\rangle_0. \tag{60}$$

我们在第 2 小节中已看到过李代数 $su(2)$ 和 $sl(2,\mathbb{R})$ 上的基灵度量的例子 (公式 (33) 和 (41)); 在那里, 这些度量被用于证明定理 24.3 和 24.4. 我们应注意, 对李

代数 $su(2)$ 的基灵度量 (33) 是欧几里得的, 而对 $sl(2,\mathbb{R})$ 的基灵度量则是伪欧几里得的 ((1,2) 型).

设 $\mathfrak{g}$ 为变换群 G 的李代数; 又设在 $\mathfrak{g}$ 上给出了基灵度量 $\langle,\rangle_0$. 利用在上一小节中构建的左不变场, 可以构造整个群上的度量 $\langle,\rangle$:

$$\langle L_X, L_Y\rangle = \langle X, Y\rangle. \tag{61}$$

因此, 两个左不变向量场的内积被认为恒等于它们在群的单位元所取值的内积. 由于引理 24.3, 这个约定完全确定了 G 上的一个度量.

定义 24.6. 由等式 (61) 定义的度量被称做群 G 的*基灵度量*.

例 24.9. 设 $G = SO(n,\mathbb{R})$. 我们指出, 这个群的基灵度量可以由所有矩阵的空间 $\mathbb{R}^{n^2}$ 上的欧氏度量诱导. 这个欧氏度量为

$$\langle X, Y\rangle = \sum_{i,j} x^i_j y^i_j, \quad X = (x^i_j), \quad Y = (y^i_j). \tag{62}$$

我们注意到, 这个表达式可以转换为

$$\langle X, Y\rangle = \mathrm{Tr}(XY^{\mathrm{T}}), \tag{63}$$

其中 Tr 为矩阵的迹; $\mathrm{Tr}\,A = a^i_i$, 其中 $A = (a^i_j)$. 曲面 $SO(n,\mathbb{R})$ 完全位于半径为 $\sqrt{n}$ 的球面上, 这是因为对 n 阶正交矩阵 A 有

$$\langle A, A\rangle = \mathrm{Tr}(AA^{\mathrm{T}}) = \mathrm{Tr}(1) = n.$$

设 $X, Y \in so(n,\mathbb{R}), A \in SO(n,\mathbb{R})$. 我们要证明, 在由欧氏度量 (62) 在曲面 $SO(n,\mathbb{R})$ 上诱导的度量下成立等式

$$\langle L_X(A), L_Y(A)\rangle = \mathrm{Tr}(XY^{\mathrm{T}}) = \langle L_X(1), L_Y(1)\rangle. \tag{64}$$

由诱导度量的定义,

$$\begin{aligned}\langle L_X(A), L_Y(A)\rangle &= \mathrm{Tr}(AX(AY)^{\mathrm{T}}) = \mathrm{Tr}(AXY^{\mathrm{T}}A^{\mathrm{T}})\\ &= \mathrm{Tr}(A^{\mathrm{T}}AXY^{\mathrm{T}}) = \mathrm{Tr}(XY^{\mathrm{T}})\end{aligned}$$

(我们用了乘积因子的循环置换不改变乘积的迹这个事实以及正交条件 $A^{\mathrm{T}}A = 1$).

还需证明 $\mathrm{ad}\,X$ 对 $X \in so(n,\mathbb{R})$ 时在此度量下为反称. 如果 $Y, Z \in so(n,\mathbb{R})$, 即 $Y^{\mathrm{T}} = -Y, Z^{\mathrm{T}} = -Z$, 则

$$\langle Y, Z\rangle = \mathrm{Tr}(YZ^{\mathrm{T}}) = -\mathrm{Tr}(YZ).$$

因此,

$$\langle \mathrm{ad}\,X(Y), Z\rangle = -\mathrm{Tr}(XYZ) + \mathrm{Tr}(YXZ),$$

$$\langle Y, \text{ad } X(Z)\rangle = -\text{Tr}(YXZ) + \text{Tr}(YZX).$$

这两个表达式仅仅差一个符号. 故而我们得到了群 $SO(n,\mathbb{R})$ 的基灵度量 (从而在它的任意子群上), 它是由欧氏空间 $\mathbb{R}^{n^2}$ 上的欧氏度量诱导的.

如果 G 是作为酉群 $U(n)$ 的子群实现的, 则可以利用实化运算, 给出嵌入 $U(n) \subset SO(2n,\mathbb{R})$. 从而可以直接写出群 $U(n)$ 的基灵度量

$$\langle X, Y\rangle = \text{Re Tr}(X\bar{Y}^{\text{T}}) = -\text{Re Tr}(XY), \quad X, Y \in u(n).$$

6. 三维李代数的分类

设给出了三维李代数 L; 又设 e_1, e_2, e_3 为它的一组基.L 中的交换规则给出了"结构常数" 张量 c_{ij}^k, 它们由下式定义

$$[e_i, e_j] = c_{ij}^k e_k, \quad i, j = 1, 2, 3. \tag{65}$$

显然

$$c_{ij}^k = -c_{ji}^k. \tag{66}$$

由雅可比恒等式 (见前面) 得到关系

$$c_{ij}^k c_{kl}^m + c_{jl}^k c_{ki}^m + c_{li}^k c_{kj}^m = 0. \tag{67}$$

通过换基, 我们要把 c_{ij}^k 转化为最简单的形式. 为此, 我们通过对称张量 (b^{ij}) 的分量和向量 (a_i) 的分量来表示它: 令

$$c_{ij}^k = \varepsilon_{ijl} b^{lk} + \delta_j^k a_i - \delta_i^k a_j. \tag{68}$$

(这是 9 个线性方程的方程组, 它有 9 个未知量 $a_1, a_2, a_3, b^{kl} = b^{lk}, k, l = 1, 2, 3$, 总有解.) 反称条件已经满足, 而雅可比恒等式 (67) 给出

$$b^{ij} a_j = 0, \tag{69}$$

即向量 (a_i) 或者为零, 或者为张量 (b^{ij}) 的零特征值所对应的特征向量. 我们化张量 (b^{ij}) 为对角型: $b^{ij} = b^{(i)}\delta^{ij}$, 其中 $b^{(1)}, b^{(2)}, b^{(3)}$ 为特征值. 因为 (a_i) 为特征向量, 可以设 $(a_i) = (a, 0, 0)$. 由 (69) 得到 $b^{(1)}a = 0$; 即或 $b^{(1)}$ 为零或 a 为零. 交换规则 (65) 此时化为

$$[e_1, e_2] = ae_2 + b^{(3)}e_3, \quad [e_2, e_3] = b^{(1)}e_1, \quad [e_3, e_1] = b^{(2)}e_2 - ae_3.$$

还需要决定向量 e_1, e_2, e_3 的符号变化, 以及进行乘以任意的正常数. 考虑到了这些可能性后, 我们得到了下面三维李代数的表格 (比安基分类):

类型	a	$b^{(1)}$	$b^{(2)}$	$b^{(3)}$	类型	a	$b^{(1)}$	$b^{(2)}$	$b^{(3)}$
Ⅰ	0	0	0	0	Ⅴ	1	0	0	0
Ⅱ	0	1	0	0	Ⅳ	1	0	0	1
VII_0	0	1	1	0	Ⅶ	a	0	1	1
VI_0	0	1	-1	0	Ⅲ$(a=1)$				
Ⅸ	0	1	1	1	Ⅵ$(a\neq 1)$	a	0	1	-1
Ⅷ	0	1	1	-1					

类型 Ⅰ 为阿贝尔李代数 (平移群的李代数). 类型 Ⅸ是 $SO(3)$ 群的李代数.

7. 共形群的李代数

考虑对应于欧氏和伪欧氏空间的共形变换所对应的向量场. 像在 §15 中证明的, 维数 ⩾3 的共形变换非常少 (参看定理 15.2). 它们对应于空间 $\mathbb{R}^n_{p,q}$ 的运动 (伪旋转及平移), 伸缩, 对某个中心的反演. 我们引进下面的场, 它们由一阶微分算子所描述:

$$1)\ \Omega_{ab}=g_{ac}x^c\frac{\partial}{\partial x^b}-g_{bc}x^c\frac{\partial}{\partial x^a},\quad a,b=1,\cdots,n\ (\text{伪旋转}),$$

$$2)\ P_a=\frac{\partial}{\partial x^a}\ (\text{平移}),$$

$$3)\ D=x^a\frac{\partial}{\partial x^a}(\text{伸缩}),\tag{70}$$

$$4)\ K_a=2g_{ac}x^cx^b\frac{\partial}{\partial x^b}-g_{bc}x^bx^c\frac{\partial}{\partial x^a}\ (\text{反演}).$$

对于欧氏度量我们有 $g_{ac}=\delta_{ac}$, 从而变换 $\exp(t\Omega_{ab})$ 给出了平面 (x^a,x^b) 中的旋转; 对于伪欧氏度量 $g_{ac}=\lambda_a\delta_{ac},\lambda_a=\pm1$. 这时我们得到, 变换 $\exp(t\Omega_{ab})$ 或者给出了旋转 (如果 $\lambda_a=\lambda_b$), 或者给出了平面 (x^a,x^b) 中的洛伦兹变换 (如果 $\lambda_a=-\lambda_b$). 变换 $\exp\left(t\dfrac{\partial}{\partial x^a}\right)$ 是沿 x^a 轴的平移, 而 $\exp\left(tx^a\dfrac{\partial}{\partial x^a}\right)$ 则是伸缩 $D(x)=tx$. 求变换 $\exp(tK_a)$ 的单参数群相当复杂, 这是因为场 K_a 不是线性的. 稍后我们要证明所有这些变换在 $\mathbb{R}^n_{p,q}$ 上是共形的.

注 另外, 靠近 $\mathbb{R}^n$ 或球面 S^n 的恒同变换的任意共形变换可以表示为 $\exp(tA)$ 的形式, 其中

$$A=\sum_{a,b=1}^n\lambda^{ab}\Omega_{ab}+\sum_{a=1}^n\mu^aP_a+\gamma D+\sum_{a=1}^n\delta^aK_a.$$

上面所写出的这些向量场 (70) 构成了李代数. 不难验证其交换关系为

$$
\begin{aligned}
&[\Omega_{ab}, \Omega_{cd}] = g_{ac}\Omega_{bd} - g_{bc}\Omega_{ad} + g_{ad}\Omega_{cb} - g_{bd}\Omega_{ca};\\
&[\Omega_{ab}, P_c] = g_{ac}P_b - g_{bc}P_a; \quad [\Omega_{ab}, K_c] = g_{ac}K_b - g_{bc}K_a;\\
&[\Omega_{ab}, D] = [P_a, P_b] = [K_a, K_b] = 0;\\
&[P_a, K_b] = 2(g_{ab}D + \Omega_{ab}); \quad [P_a, D] = P_a; \quad [K_a, D] = -K_a.
\end{aligned}
\tag{71}
$$

现设度量 g_{ab} 为欧氏的, 即 $g_{ab} = \delta_{ab}$. 考虑对应于伪旋转群 $SO(n+1,1)$ 的李代数, 由相似的那些向量场 $\Omega_{\mu\nu}, \mu, \nu = 1, \cdots, n+2$ 给出. 建立下面的对应关系

$$
\begin{aligned}
&\Omega_{ab} \mapsto \Omega_{\mu,\nu}, \quad \mu = a = 1, \cdots, n, \quad \nu = b = 1, \cdots, n,\\
&P_a \mapsto \Omega_{a,n+1} - \Omega_{a,n+2},\\
&K_a \mapsto \Omega_{a,n+1} + \Omega_{a,n+2},\\
&D \mapsto \Omega_{n+1,n+2}.
\end{aligned}
\tag{72}
$$

可直接验证下面重要的事实.

命题 24.1. 对应关系 (72) 是李代数间的同构.

因此欧氏情形的李代数 (71) 同构于对应伪旋转的群 $SO(n+1,1)$ 的李代数 (即 A 为李代数 (71) 中元的变换 $\exp(tA)$ 为伪旋转).

注 1 虽然在 $n = 2$ 时有许多局部的共形变换, 但我们挑出一个球面 $S^2 \to S^2$ 的分式线性变换的子群

$$z \mapsto \frac{az+b}{cz+d}, \quad ad - bc \neq 0,$$

我们曾经用过它的一些子群来研究欧氏平面 $\mathbb{R}^2$ 的运动, 罗巴切夫斯基平面 L^2 的及球面 S^2 的运动 (见 §9, 10, 13). 这个群同构于 $SL(2,\mathbb{C})/\pm 1$(见 §13), 而且经过球极投影到 $\mathbb{R}^2$ 后恰好由 $\mathbb{R}^2$ 的旋转, 平移, 伸缩和反演生成.

同构 (72) 给出了所有迹为零的复 2×2 矩阵的李代数与洛伦兹群 $SO(3,1)$ 的李代数的同构, 后者是作为形如 $\exp(tA)$ 的伪旋转群实现的. 称这个同构为复 2×2 矩阵的洛伦兹群 $SO(3,1)$ 的"半旋量表示". 也存在复共轭表示.

注 2 伪欧氏空间 $\mathbb{R}^n_{p,q}$ 的共形变换群的李代数同构于 $so(p+1, q+1)$(请验证!).

最后证明下面命题.

定理 24.8. 设 $g_{ab} = \delta_{ab}$ 且 $A = \lambda^{ab}\Omega_{ab} + \mu^a P_a + \gamma D + \delta^a K_a$. 于是形如 $S_t = \exp(tA)$ 的变换是欧氏空间 $\mathbb{R}^n$ 的 (局部) 共形变换.

要证明此定理需要考察向量场 A 并证明由它定义的 (局部) 变换群 $S_t = \exp(tA)$ 是共形的 (如果我们希望要谈的不是局部而是整体的变换, 则应该考虑球面 S^n 上的这个场, 它是由逆于球极投影的变换从 $\mathbb{R}^n$ 映到球面上, 它把欧氏坐标共形地引进到 S^n 上. 但是这并不存在). 如果向量场 (u^a) 给出了欧氏度量的运动, 则它应该满足条件

$$\frac{\partial u^a}{\partial x^b} + \frac{\partial u^b}{\partial x^a} = 0.$$

这表明量度畸变距离的“形变张量”化为了零 (参看 §23). 对给出了欧氏度量的共形变换的向量场 (u^a) 来说, 形变张量应该与这个度量成比例:

$$\frac{\partial u^a}{\partial x^b}+\frac{\partial u^b}{\partial x^a}=\gamma(x)\delta_{ab}, \tag{73}$$

其中 $\gamma(x)$ 为光滑函数. 在这种情形下, 如果 $A=u^a\dfrac{\partial}{\partial x^a}$, 我们设得到了变换

$$S_t=\exp(tA),$$

从而度量的变换为

$$S_t^*(g_{ab})_{ij}=S_t^*(\delta_{ab})_{ij}=[1+t\mu(x)]\delta_{ij}+O(t^2).$$

(事实上, 由定义 (参看 §22.1)

$$S_t^*(\delta_{ab})_{ij}=\delta_{ab}\frac{\partial x^a}{\partial y^i}\frac{\partial x^b}{\partial y^j}=\sum_{a=1}^{n}\frac{\partial x^a}{\partial y^i}\frac{\partial x^a}{\partial y^j}.$$

由公式 (23.1.6) 得到 $y^i=x^i+tu^i(x^1,\cdots,x^n)=o(t)$, 于是

$$\frac{\partial y^i}{\partial x^a}=\delta_a^i+t\frac{\partial u^i}{\partial x^a}+o(t).$$

由此等式连同 (73) 表明

$$\sum_{i=1}^{n}\frac{\partial y^i}{\partial x^a}\frac{\partial y^i}{\partial x^b}=[1+t\gamma(x)]\delta_{ab}+O(t^2).$$

逆转这个最后面的等式并代入 $S_t^*(\delta_{ab})$ 的表达式便有了所想要的结果.)

直接的验证使我们确信整个场 (72) 都满足条件 (73).(例如, 令 $\partial_u\left(=u^i\dfrac{\partial}{\partial x^i}\right)=K_a$, 则得到

$$u^j=2x^ax^j,\quad j\neq a;\quad u^a=2(x^a)^2-\sum_i(x^i)^2.$$

由此立即得出了 (73))

剩下来要证明的是, 反过来由条件 (73) 得出由向量场

$$\dot{x}^a=u^a(x^1,\cdots,x^n) \tag{74}$$

定义的变换 S_t 或者等价于说变换 $\exp(tA)$ 的共形性. 证明如下 (它类似于对于运动的相应证明). 对有限时间段 Δt 的位移可以分解为一系列 (相邻接的) 时间段 $\dfrac{\Delta t}{N}$ 的位移:

$$S_{\Delta t}=S_{(\Delta t/N)}\circ\cdots\circ S_{(\Delta t/N)}.$$

变换 S_τ 的线性部分当 $\tau - \Delta t/N$ 较小时有形式

$$S_\tau = 1 + \tau A + O(\tau^2), \quad A = \left(\frac{\partial u^a}{\partial x^b}\right).$$

由于 $\Delta t/N$ 较小, 对在点 $x_0 = (x_0^1, \cdots, x_0^n)$ 的向量 ξ 有

$$\begin{aligned}
&\xi_0^a = \xi^a \to \xi^a + \left(\frac{\Delta t}{N}\right)\frac{\partial u^a}{\partial x^b}(x_0)\xi^b + O\left(\frac{\Delta t^2}{N^2}\right) = \xi_1^a,\\
&\xi_1^a \to \xi_2^a = \xi_1^a + \left(\frac{\Delta t}{N}\right)\frac{\partial u^a}{\partial x^b}(x_1)\xi_1^b + O\left(\frac{\Delta t^2}{N^2}\right),\\
&\cdots\cdots\\
&\xi_{N-1}^a \to \xi_N^a = \xi_{N-1}^a + \left(\frac{\Delta t}{N}\right)\frac{\partial u^a}{\partial x^b}(x_{N-1})\xi_{N-1}^b + O\left(\frac{\Delta t^2}{N^2}\right).
\end{aligned} \tag{75}$$

点 $x_q, q = 0, \cdots, N-1$, 位于方程 (74) 的积分轨线 $x^a(t)$ 上, 使得

$$\begin{aligned}
&x^a(0) = x_0^a,\\
&x^a\left(\frac{\Delta t}{N}\right) = x_1^a,\\
&\cdots\cdots\\
&x^a\left((N-1)\frac{\Delta t}{N}\right) = x_{N-1}^a.
\end{aligned}$$

对于向量 ξ_i 的自身内积, 我们由 (75) 得到

$$\begin{aligned}
&\langle\xi_1, \xi_1\rangle = \left(1 + \mu(x_0)\frac{\Delta t}{N}\right)\langle\xi_0, \xi_0\rangle + O\left(\frac{\Delta t^2}{N^2}\right),\\
&\langle\xi_2, \xi_2\rangle = \left(1 + \mu(x_1)\frac{\Delta t}{N}\right)\langle\xi_1, \xi_1\rangle + O\left(\frac{\Delta t^2}{N^2}\right),\\
&\cdots\cdots\\
&\langle\xi_N, \xi_N\rangle = \left(1 + \mu(x_{N-1})\frac{\Delta t}{N}\right)\langle\xi_{N-1}, \xi_{N-1}\rangle + O\left(\frac{\Delta t^2}{N^2}\right).
\end{aligned}$$

因为全部步数为 N, 故

$$\langle\xi_N, \xi_N\rangle = \left[1 + \left(\sum_{i=0}^{N-1}\mu(x_i)\right)\frac{\Delta t}{N}\right]\langle\xi_0, \xi_0\rangle + O\left(N \cdot \frac{\Delta t^2}{N^2}\right). \tag{76}$$

当固定的 $\Delta t =$ 常数及 $N \to \infty$ 时, 这个关系式的右端趋向于

$$\left[1 + \int_{\Delta t}\mu(x(t))dt\right]\langle\xi_0, \xi_0\rangle - \rho(x_0)\langle\xi_0, \xi_0\rangle,$$

从而我们得到了变换 $S_{\Delta t}$ 的共形性. 由此得到了我们的论断.

注

1. 由此定理的证明得到, 如果形变场的张量为零, 则 $S_t = \exp(tA)$ 为运动, 其中 $A = u^a \frac{\partial}{\partial x^a}$. 这完全等同于说, 所有线性变换 $S_t^{-1}\frac{dS_t}{dt}$ 在每个点上为反称的:

$$S_t^{-1}\frac{dS_t}{dt} = B(t) - \left(\frac{\partial u^a}{\partial x^b}\right); \quad B = \frac{1}{2}\left(\frac{\partial u^a}{\partial x^b} - \frac{\partial u^b}{\partial x^a}\right) = \left(\frac{\partial u^a}{\partial x^b}\right).$$

2. 另外, 可能出现这种情形, 即变换 $B(x) = \left(\frac{\partial u^a}{\partial x^b}\right)$ 在所有点 $x \in \mathbb{R}^n$ 生成了整个 $n \times n$ 矩阵的李代数中的子代数. 例如, 对偶数 n, 它可以是所有复的 $\frac{n}{2} \times \frac{n}{2}$ 矩阵的代数.

习题

24.1. 证明在上面最后一种情形中, 变换 S_t 给出了 $\mathbb{C}^{n/2} \to \mathbb{C}^{n/2}$ 的全纯变换. 对于 $n/2 > 1$, 一般来说, 这个变换不是共形的.

24.2. 证明, 如果所有 $B(x)$ 具有零迹, 则所有变换 S_t 保持体积不变.

24.3. 证明下列李代数间的同构:

a) $su(1,1) \simeq sl(2,\mathbb{R})$,

b) $su(2) \times su(2) \simeq so(4)$,

c) $sl(2,\mathbb{C}) \simeq so(1,3)$,

d) $so(1,2) \simeq \mathbb{R}^3_{1,2}$ 中向量关于 "向量积" 的代数 (参看习题 6.1).

24.4. 计算在么模四元数群上的左不变向量场.

24.5. 证明在群 $SO(n)$ 上的基灵度量可以写为

$$dl^2(g) = \mathrm{Tr}(g^{-1}dg \cdot g^{-1}dg).$$

24.6. 设 g^0_{ij} 为李代数 $\mathfrak{g}$ 上的基灵度量, 它是在基 $X_1, \cdots, X_n$ 下表示的. 设 $[X_i, X_j] = c^k_{ij}X_k, c_{kij} = g^0_{kl}c^l_{ij}$. 证明张量 c_{ijk} 对所有三个指标都是反对称的.

24.7 证明对群 $SO(n,\mathbb{R})$, 内自同构 $B \mapsto ABA^{-1}$ 为基灵度量的运动.

24.8. 对于群 $SO(p,q)$, 可以通过伪欧氏度量 $\langle X, Y\rangle = \mathrm{Tr}(GXGY^{\mathrm{T}})$ 的限制得出基灵度量 (其中 G 为 (p,q) 型的度量的矩阵). 求所得到的伪黎曼度量的类型.

24.9. 证明在上面描述过的所有基灵度量都可以作为 $\mathrm{Tr}(\mathrm{ad}\, X, \mathrm{ad}\, Y)$ 得到 (精确到常数乘积因子).

24.10. 群 $SL(2,\mathbb{R})/\{\pm 1\}$ 的作用同于罗巴切夫斯基度量的运动群 (参见 §13.2). $SL(2,\mathbb{R})$ 中每个单参数子群 (见习题 14.5) 对应于罗巴切夫斯基平面的单参数微分同胚群. 求在罗巴切夫斯基平面上 (克莱因模型) 的相应的向量场. 计算它们的换位子.

24.11. 计算 $\mathbb{R}^n$ 中的仿射群的李代数; n 维欧氏空间运动群的李代数; 闵可夫斯基空间 $\mathbb{R}^3_{1,3}$ 运动群的李代数. 写出在 $\mathbb{R}^n$ 和 $\mathbb{R}^3_{1,2}$ 中相应的生成元.

24.12. 证明 $\mathbb{R}^3$ 中对应于群 $SO(3,\mathbb{R})$ 的作用的线性向量场 L_X, L_Y, L_Z, 它们切于中心在坐标原点的任一球面. 求出在球面坐标下, 单位球上的相对应的一阶微分算子.

24.13. 定义在群 G 上的右不变向量场为形如 $R_X(A) = -XA$ 的向量场在此群上的限制. 证明 $[R_X, R_Y] = R_{[X,Y]}, [L_X, R_Y] \equiv 0$.

24.14. 解释清楚李代数 $so(1,2)$ 和 $sl(2,\mathbb{R})$ 属于在第 6 小节中所描述的那些类型中的哪一类.

第四章　张量的微分学

§25. 反称张量的微分

1. 反称张量的梯度

大多数的物理定律都是以物理量之间的微分关系描述的. 这些物理量中许多量都是空间或空间中区域上的张量场 (特别是向量场). 因此我们感兴趣的问题是, 在张量上的微分运算一般应是什么样子, 而且它还应是在一种明确的意义下与坐标的选取无关. (我们将在后面阐明是什么样的意义). 例如, 有这样的最简单运算: 如果函数 $f(x,\alpha)$ 或张量场 $T^{i_1\cdots i_p}_{j_1\cdots j_q}(x,\alpha)$ 依赖于空间中的点 $x=(x^1,x^2,x^3)$ 和某个参数 α, 这个参数与空间无关, 那么可以在每个给定点上取关于参数 α 的偏导数 $\dfrac{\partial f}{\partial\alpha}(x,\alpha)$ 或 $\dfrac{\partial T^{i_1\cdots i_p}_{j_1\cdots j_q}(x,\alpha)}{\partial\alpha}$. 在经典力学中这样的参数可以是时间 $t=\alpha$. 这个运算与空间 (x^1,x^2,x^3) 的几何无关, 并且求导是孤立地在每个点上进行. 另一个所熟知的与黎曼度量无关的微分运算是取函数 (标量场) 的梯度

$$\left(\frac{\partial f}{\partial x^1},\frac{\partial f}{\partial x^2},\frac{\partial f}{\partial x^3}\right)=\operatorname{grad} f.$$

这是一个余向量, 说它由函数 f 以不变的方式构造的意思是表明在坐标变换下, 它的数值记号按张量的规则变化, 即

$$x=x(z),\quad \frac{\partial f}{\partial z^j}=\frac{\partial f}{\partial x^i}\frac{\partial x^i}{\partial z^j}.$$

常常也遇到下面的反称张量的高维情形梯度的推广.

定义 25.1. 设 $T_{i_1\cdots i_k}$ 为在具坐标 $x^1,\cdots,x^n$ 的 n 维空间中的张量, 它对所有下指标为反对称的, 其中 $i_q=1,\cdots,n$. 它的梯度是指 $(0,k+1)$ 型的一个反称张量, 其分量为

$$(dT)_{j_1\cdots j_{k+1}}=\sum_{q=1}^{k+1}(-1)^{q+1}\frac{\partial T_{j_1\cdots\widehat{j}_q\cdots j_{k+1}}}{\partial x^{j_q}}. \tag{1}$$

(这里的记号 $\widehat{j}_k$ 表示指标 j_q 已被略去.)

在验证 dT 为张量之前, 先考虑几个例子.

1) 如 $k+1=1$, 从而 $T=f(x)$ 为函数, 故由定义 $(dT)_i=\dfrac{\partial T}{\partial x^i}$; 因而这是通常的梯度.

2) 如果 $T=(T_i)$ 为余向量, 则 $(dT)_{ij}=\dfrac{\partial T_j}{\partial x^i}-\dfrac{\partial T_i}{\partial x^j}=-(dT)_{ji}$. 常常称此余向量场为旋度. 我们以记号 rot 表示它. 旋度是 (0,2) 型的反称张量.

注 如果在 $n=3$ 的欧氏空间及坐标 x^1,x^2,x^3 中, 则张量 $(dT)_{ij}$ 通常与向量 $(\eta^k)=\operatorname{rot}T=*(dT)$ (见 §19.3) 重合, 其中 $\eta^1=\dfrac{\partial T_3}{\partial x^2}-\dfrac{\partial T_2}{\partial x^3}=(dT)_{23},\eta^2=\dfrac{\partial T_1}{\partial x^3}-\dfrac{\partial T_3}{\partial x^1}=(dT)_{31}=-(dT)_{13},\eta^3=\dfrac{\partial T_2}{\partial x^1}-\dfrac{\partial T_1}{\partial x^2}=(dT)_{12}$ 或者表示为 $\eta^i=\dfrac{1}{2}\varepsilon^{ijk}(dT)_{jk}$.

3) 如果 $n=3$ 且给出了反称张量 $T_{ij}=-T_{ji}$, 那么 3 阶反称张量 dT 的形式为

$$(dT)_{123}=\frac{\partial T_{12}}{\partial x^3}-\frac{\partial T_{13}}{\partial x^2}+\frac{\partial T_{23}}{\partial x^1},\quad (dT)_{ijk}=\varepsilon_{ijk}(dT)_{123}.$$

注 如果坐标 x^1,x^2,x^3 为欧氏的, 从而按前面所给出的规则, 反称张量所对应的相伴向量为 $\eta^1=T_{23},\eta^2=-T_{13},\eta^3=T_{12};(\eta^i=\dfrac{1}{2}\varepsilon^{ijk}T_{jk})$, 于是

$$(dT)_{123}=\frac{\partial\eta^1}{\partial x^1}+\frac{\partial\eta^2}{\partial x^2}+\frac{\partial\eta^3}{\partial x^3}=\frac{\partial\eta^i}{\partial x^i}.$$

对在欧氏坐标下向量场 $(\eta^i)=\eta$ 赋予数值 $\dfrac{\partial\eta^i}{\partial x^i}=\operatorname{div}\eta$ 的运算被称做发散量.

我们转向验证前面定义的正确性.

定理 25.1. $(0,k)$ 型的 k 阶反称张量的梯度 dT 是个 $(0,k+1)$ 型的 $k+1$ 阶反称张量.

证明 (只证 $k=0,1$). 设给出变换

$$x^i=x^i(x^{1'},\cdots,x^{n'}),\quad i=1,\cdots,n. \tag{2}$$

由定义, 在任意坐标系下均有

$$(dT)_{i_1\cdots i_{k+1}}=\sum_q(-1)^{q+1}\frac{\partial T_{i_1\cdots\widehat{i}_q\cdots i_{k+1}}}{\partial x^{i_q}}. \tag{3}$$

设 $T_{i_1\cdots i_k}$ 为在坐标 (x) 下的张量的分量, $T_{i'_1\cdots i'_k}$ 为在坐标 (x') 下的张量. 由定义

$$T_{i'_1\cdots i'_k}=T_{i_1\cdots i_k}\frac{\partial x^{i_1}}{\partial x^{i'_1}}\cdots\frac{\partial x^{i_k}}{\partial x^{i'_k}}. \tag{4}$$

另外, 由张量的梯度定义有

$$(dT)_{i'_1\cdots i'_{k+1}}=\sum_q(-1)^{q+1}\frac{\partial T_{i'_1\cdots\widehat{i'_q}\cdots i'_{k+1}}}{\partial x^{i'_q}}, \tag{5}$$

$$(dT)_{i_1\cdots i_{k+1}}=\sum_p(-1)^{p+1}\frac{\partial T_{i_1\cdots\hat{i}_p\cdots i_{k+1}}}{\partial x^{i_p}}. \tag{6}$$

为了证明定理, 借助于公式 (4),(5) 和 (6) 必须证实在变换之后, 梯度 $dT_{(i')}$ 可按照张量的规则通过 $dT_{(i)}$ 表达. 由于计算的繁琐, 我们只对 $k=1,k+1=2$ 的情形进行完整的证明 ($k=0$ 情形时定理论断的正确性已在 §16 中证明过).

设 T_i 为余向量; 于是

$$(dT)_{ij}=\frac{\partial T_j}{\partial x^i}-\frac{\partial T_i}{\partial x^j},\quad T_{i'}=T_i\frac{\partial x^i}{\partial x^{i'}},$$

我们有

$$\begin{aligned}(dT)_{l'k'}&=\frac{\partial T_{k'}}{\partial x^{l'}}-\frac{\partial T_{l'}}{\partial x^{k'}}=\frac{\partial}{\partial x^{l'}}\left(T_i\frac{\partial x^i}{\partial x^{k'}}\right)-\frac{\partial}{\partial x^{k'}}\left(T_j\frac{\partial x^j}{\partial x^{l'}}\right)\\&=\frac{\partial T_i}{\partial x^{l'}}\frac{\partial x^i}{\partial x^{k'}}+T_i\frac{\partial^2x^i}{\partial x^{l'}\partial x^{k'}}-\frac{\partial T_j}{\partial x^{k'}}\frac{\partial x^j}{\partial x^{l'}}-T_j\frac{\partial^2x^j}{\partial x^{k'}\partial x^{l'}}\\&=\left(\frac{\partial T_i}{\partial x^j}\frac{\partial x^j}{\partial x^{l'}}\right)\frac{\partial x^i}{\partial x^{k'}}-\left(\frac{\partial T_j}{\partial x^i}\frac{\partial x^i}{\partial x^{k'}}\right)\frac{\partial x^j}{\partial x^{l'}}\\&=\left(\frac{\partial T_i}{\partial x^j}-\frac{\partial T_j}{\partial x^i}\right)\frac{\partial x^i}{\partial x^{k'}}\frac{\partial x^j}{\partial x^{l'}}=(dT)_{ji}\frac{\partial x^i}{\partial x^{k'}}\frac{\partial x^j}{\partial x^{l'}}.\end{aligned}$$

因此在此情形下的定理已证完. □

2. 形式的外微分

我们给出反称张量的梯度运算的另一个定义, 这个运算利用了它与微分形式的关联. 张量 $T_{i_1\cdots i_k}$ 对应于微分形式

$$\omega=\sum_{i_1<\cdots<i_k}T_{i_1\cdots i_k}dx^{i_1}\wedge\cdots\wedge dx^{i_k}. \tag{7}$$

我们定义 $k+1$ 阶形式 $d\omega$ 为

$$d\omega=\sum_{\substack{i_0\\ i_1<\cdots<i_k}}\frac{\partial T_{i_1\cdots i_k}}{\partial x^{i_0}}dx^{i_0}\wedge dx^{i_1}\wedge\cdots\wedge dx^{i_k}. \tag{8}$$

特别, 如果 $\omega=f$ (标量), 则 $d\omega=df=\dfrac{\partial f}{\partial x^i}dx^i$ 为函数的微分.

定理 25.2. 成立恒等式:

$$d\omega = \sum_{j_1<\cdots<j_{k+1}} (dT)_{j_1\cdots j_{k+1}} dx^{j_1} \wedge \cdots \wedge dx^{j_{k+1}}.$$

证明 由 dT 的定义有

$$\begin{aligned}
&\sum_{j_1<\cdots<j_{k+1}} (dT)_{j_1\cdots j_{k+1}} dx^{j_1} \wedge \cdots \wedge dx^{j_{k+1}} \\
=& \sum_{j_1<\cdots<j_{k+1}} \sum_q (-1)^{q+1} \frac{\partial T_{j_1\cdots \hat{j}_q \cdots j_{k+1}}}{\partial x^{j_q}} dx^{j_1} \wedge \cdots \wedge dx^{j_{k+1}}.
\end{aligned}$$

我们知道有 $dx^{j_1} \wedge \cdots \wedge dx^{j_{k+1}} = (-1)^{q+1} dx^{j_q} \wedge dx^{j_1} \wedge \cdots \wedge \widehat{dx^{j_q}} \wedge \cdots \wedge dx^{j_{k+1}}$. 于是最后面的表达式可以表示为

$$\sum_q \sum_{j_1<\cdots<j_{k+1}} \frac{\partial T_{j_1\cdots \hat{j}_q \cdots j_{k+1}}}{\partial x^{j_q}} dx^{j_q} \wedge dx^{j_1} \wedge \cdots \wedge \widehat{dx^{j_q}} \wedge \cdots \wedge dx^{j_{k+1}}.$$

现在我们变换取和的指标: 令 $j_1 = i_1, \cdots, j_{q-1} = i_{q-1}, j_{q+1} = i_q, \cdots, j_{k+1} = i_k$. 显然, $i_1 < \cdots < i_k$, 而指标 j_q 跑过 (考虑所有的项) 从 1 到 n 的每个数. 从而得到了定理的论断. □

定理 25.3. 取反称张量的二重梯度运算给出了恒等式

$$d(dT) = 0 \quad 或 \quad d(d\omega) = 0. \tag{9}$$

证明 我们有

$$\begin{aligned}
\omega &= \sum_{i_1<\cdots<i_k} T_{i_1\cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}; \\
d\omega &= \sum_{\substack{p \\ i_1<\cdots<i_k}} \frac{\partial T_{i_1\cdots i_k}}{\partial x^p} dx^p \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}; \\
d(d\omega) &= \sum_{\substack{p,q \\ i_1<\cdots<i_k}} \frac{\partial^2 T_{i_1\cdots i_k}}{\partial x^q \partial x^p} dx^q \wedge dx^p \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}.
\end{aligned}$$

然而表达式 $\dfrac{\partial^2 T_{i_1\cdots i_k}}{\partial x^q \partial x^p}$ 对指标 q, p 对称, 而 $dx^q \wedge dx^p$ 则为反称:$dx^p \wedge dx^q = -dx^q \wedge dx^p$, 故而将它们并项便等于零. 证完. □

注 对形式的微分公式可以写为

$$d\left(\sum_{i_1<\cdots<i_k} T_{i_1\cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}\right)$$

$$= \sum_{i_1<\cdots<i_k} (dT_{i_1\cdots i_k}) \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}, \tag{10}$$

其中 $T_{i_1\cdots i_k}$ 被看作标量函数. 由此不难得到 dT 是张量的另一个证明: 因为函数的微分 $dT_{i_1\cdots i_k}$ 为张量, 故而它与 $dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ 的外积也是张量.

利用向量场的交换运算, 还可以得到微分形式的微分的一个表达式 (嘉当公式).

定理 25.4. 设 ω 为 k 阶微分形式, $X_1, \cdots, X_{k+1}$ 为光滑向量场. 于是微分形式 $d\omega$ 在向量场 $X_1, \cdots, X_{k+1}$ 上的值可通过下面公式求出:

$$(k+1)d\omega(X_1, \cdots, X_{k+1}) = \sum_i (-1)^{i-1}\partial_{X_i}\omega(X_1, \cdots, \widehat{X}_i, \cdots, X_{k+1}) +$$

$$\sum_{i<j}(-1)^{i+j}\omega([X_i, X_j], X_1, \cdots, \hat{X}_i, \cdots, \hat{X}_j, \cdots, X_{k+1}). \tag{11}$$

证明 我们对 $k=1$ 的情形进行证明. 设 $T_i dx^i = \omega$, 于是

$$2d\omega = \left(\frac{\partial T_j}{\partial x^i} - \frac{\partial T_i}{\partial x^j}\right) dx^i \wedge dx^j.$$

形式 $d\omega$ 在向量场 $X = (X^i), Y = (Y^i)$ 的值等于

$$2d\omega(X, Y) - X^i Y^j \left(\frac{\partial T_j}{\partial x^i} - \frac{\partial T_i}{\partial x^j}\right). \tag{12}$$

当 $k=1$ 时公式 (11) 有形式

$$\begin{aligned} 2d\omega(X, Y) &= \partial_X \omega(Y) - \partial_Y \omega(X) - \omega([X, Y]) \\ &= \partial_X(T_i Y^i) - \partial_Y(T_i X^i) - T_i\left(X^k \frac{\partial Y^i}{\partial x^k} - Y^k \frac{\partial X^i}{\partial x^k}\right) \\ &= X^i Y^j \left(\frac{\partial T_j}{\partial x^i} - \frac{\partial T_i}{\partial x^j}\right). \end{aligned} \tag{13}$$

公式 (12) 和 (13) 的右端相合. 定理得证. □

在两个形式的外积上的微分是如何进行的? 成立

定理 25.5. 设 ω_1, ω_2 分别为 p, q 阶的微分形式. 于是

$$d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^p \omega_1 \wedge d\omega_2. \tag{14}$$

证明 只需对微分形式 ω_1, ω_2 为单项的情形证明定理即可:

$$\omega_1 = f dx^{i_1} \wedge \cdots \wedge dx^{i_p}, \quad \omega_2 = g dx^{j_1} \wedge \cdots \wedge dx^{j_q}.$$

于是

$$\omega_1 \wedge \omega_2 = fg dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q}.$$

故而

$$\begin{aligned} d(\omega_1\wedge\omega_2) &= \frac{\partial f}{\partial x^k} g dx^k\wedge dx^{i_1}\wedge\cdots\wedge dx^{j_q} + f\frac{\partial g}{\partial x^k}dx^k\wedge dx^{j_1}\wedge\cdots\wedge dx^{j_q} \\ &= \left(\frac{\partial f}{\partial x^k}dx^k\wedge dx^{i_1}\wedge\cdots\wedge dx^{i_p}\right)\wedge\left(g dx^{j_1}\wedge\cdots\wedge dx^{j_q}\right) + \\ &\quad (-1)^p(f dx^{i_1}\wedge\cdots\wedge dx^{i_p})\wedge\left(\frac{\partial g}{\partial x^k}dx^k\wedge\cdots\wedge dx^{j_q}\right) \\ &= d\omega_1\wedge\omega_2 + (-1)^p\omega_1\wedge d\omega_2, \end{aligned}$$

在其中我们运用了等式

$$dx^k\wedge dx^{i_1}\wedge\cdots\wedge dx^{i_p}\wedge dx^{j_1}\wedge\cdots\wedge dx^{j_q} = (-1)^p dx^{i_1}\wedge\cdots\wedge dx^{i_p}\wedge dx^k\wedge dx^{j_1}\wedge\cdots\wedge dx^{j_q}.$$

定理得证. □

在有度量 g_{ij} 的情形则可定义在形式上的另一个重要的微分运算, 它将形式的阶降低一阶, 并以 δ 记之 (即反称张量的发散量):

$$\delta = *^{-1}d*, \tag{15}$$

其中 $*$ 为在 §19.3 中定义的与度量有关的那个运算. 对 δ 算子的显式表达式我们将在 §29 给出. 算子 δ 对于具有正雅可比的变换是不变的.

转而讨论例子. 考虑三维的欧氏空间.

1) 对于标量场 $f(x)$, 微分 df 是个余向量. 在使上下指标等同时 (限于坐标的正交变换) 我们得到了向量

$$\operatorname{grad} f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right).$$

2) 设 $\omega = T_1dx^1 + T_2dx^2 + T_3dx^3$. 于是

$$\begin{aligned} d\omega &= \left(\frac{\partial T_2}{\partial x^1} - \frac{\partial T_1}{\partial x^2}\right)dx^1\wedge dx^2 + \left(\frac{\partial T_3}{\partial x^2} - \frac{\partial T_2}{\partial x^3}\right)dx^2\wedge dx^3 + \\ &\quad \left(\frac{\partial T_1}{\partial x^3} - \frac{\partial T_3}{\partial x^1}\right)dx^3\wedge dx^1 \end{aligned}$$

为 2 阶形式. 将算子 $*$ 用于形式 $d\omega$, 可以构造出余向量 (1 形式)$*d\omega$(见 §19.3):

$$*d\omega = \left(\frac{\partial T_3}{\partial x^2} - \frac{\partial T_2}{\partial x^3}\right)dx^1 + \left(\frac{\partial T_1}{\partial x^3} - \frac{\partial T_3}{\partial x^1}\right)dx^2 + \left(\frac{\partial T_2}{\partial x^1} - \frac{\partial T_1}{\partial x^2}\right)dx^3. \tag{16}$$

如果将向量和余向量等同, 则此运算将向量 dT 转变为向量 $*dT = \operatorname{rot} T$, 即为向量场的旋量.

3) 又设 $\omega = T_1dx^1 + T_2dx^2 + T_3dx^3$. 计算形式 $\delta\omega = *^{-1}d*\omega$. 这是个零阶形式(即标量), 因为 δ 将阶降了 1. 我们有

$$\begin{aligned}\omega &= T_1dx^1 + T_2dx^2 + T_3dx^3\\ &\xrightarrow{*} T_1dx^2\wedge dx^3 + T_2dx^3\wedge dx^1 + T_3dx^1\wedge dx^3\\ &\xrightarrow{d} \left(\frac{\partial T_1}{\partial x^1}+\frac{\partial T_2}{\partial x^2}+\frac{\partial T_3}{\partial x^3}\right)dx^1\wedge dx^2\wedge dx^3 \xrightarrow{*^{-1}} \frac{\partial T_1}{\partial x^1}+\frac{\partial T_2}{\partial x^2}+\frac{\partial T_3}{\partial x^3}.\end{aligned}$$

结论

$$\delta(T) = \mathrm{div}\ T = \frac{\partial T_1}{\partial x^1}+\frac{\partial T_2}{\partial x^2}+\frac{\partial T_3}{\partial x^3}. \tag{17}$$

这个发散量的形式只在欧氏坐标下成立.

4) 考虑四维时空, 其坐标为 $x^0 = ct, x^1, x^2, x^3$ (其中 c 为光速), 而伪欧氏度量为

$$dl^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2$$

或者

$$g_{ij} = \begin{pmatrix}1 & & & 0\\ & -1 & & \\ & & -1 & \\ 0 & & & -1\end{pmatrix}.$$

我们记得在 §21 中讨论过, 电磁场为 2 阶反称张量 $F_{ij}, i,j = 0,1,2,3$. 由电磁场理论知道, 张量 F_{ij} 应满足麦克斯韦方程. 第一组或常说的第一对麦克斯韦方程为

$$(dF)_{ijk} = \frac{\partial F_{jk}}{\partial x^i} - \frac{\partial F_{ik}}{\partial x^j} + \frac{\partial F_{ij}}{\partial x^k} = 0 \tag{18}$$

或者简短地写为 $dF = 0$, 其中 $F = \sum\limits_{i<j} F_{ij}dx^i\wedge dx^j$. 另一种方式是利用在 §21 中引记的记号 $\mathbf{E} = (E_\alpha), E_\alpha = F_{0\alpha}; \mathbf{H} = (H^\alpha), -H^1 = F_{23}, H^2 = +F_{13}, -H^3 = F_{12}$. 对 $(i,j,k) = (1,2,3)$ 的方程 (18) 在这些记号下可具形式

$$\frac{\partial F_{12}}{\partial x^3} - \frac{\partial F_{13}}{\partial x^2} + \frac{\partial F_{23}}{\partial x^1} = 0, \quad 或 \quad \mathrm{div}\ \mathbf{H} = 0. \tag{19}$$

而余下的其他方程全体表明

$$\mathrm{rot}\ \mathbf{E} + \frac{\partial \mathbf{H}}{\partial x^0} = 0 \quad 或 \quad \mathrm{rot}\ \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t}. \tag{20}$$

因此, 方程组 (18) 等价于两个方程的组: 标量方程 (19) 和向量方程 (20), 这正是称为方程对的原因. 这些方程和伪欧氏度量没有什么联系.

麦克斯韦方程的第二对则已经需要伪欧氏度量来描述了. 它的总体形式为

$$\delta F = *d*F = -\frac{4\pi}{c}j_{(4)}, \tag{21}$$

其中 $j_{(4)}$ 为四维电流向量 $j_{(4)} = (\rho c, -\rho v^1, -\rho v^2, -\rho v^3) = (\rho, -\mathbf{j}), \rho$ 为三维空间中电荷密度. 在 §21 中得到的在伪欧氏坐标下算子 $*$ 的公式使我们能重写方程 (21) 为方程组 ("对") 为

$$\mathrm{div}\ \mathbf{E} = 4\pi\rho, \tag{22}$$

$$\mathrm{rot}\ \mathbf{H} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c}\mathbf{j}, \tag{23}$$

又是由一个标量方程和一个向量方程组成.

习题

25.1. 设 $X_1, \cdots, X_n$ 为在 n 维区域中每点均线性无关的向量场,$\omega^1, \cdots, \omega^n$ 为 1 形式的对偶基: $\omega^i(X_j) = \delta^i_j$. 证明

$$d\omega^k = \frac{1}{2}c^k_{ij}\omega^i \wedge \omega^j,$$

其中 c^k_{ij} 的定义为

$$[X_i, X_j] = c^k_{ij}X_k.$$

25.2. 证明对光滑映射 F 和任意的形式 ω 成立等式 $F^*(d\omega) = dF^*(\omega)$ (见 §22). 由此有 $L_\xi(d\omega) = d(L_\xi\omega)$.

25.3. 对每个向量场 $X = (X^j)$ 我们引进在形式上的算子 $i(X)$:

$$[i(X)\omega]_{j_1\cdots j_{k-1}} = X^j\omega_{jj_1\cdots j_{k-1}};$$

其中 $\omega = (\omega_{j_1\cdots j_k})$ 为 k 阶形式.

a) 证明 $i(X)$ 为反微分:

$$i(X)(\omega_1 \wedge \omega_2) = (i(X)\omega_1) \wedge \omega_2 + (-1)^k\omega_1 \wedge (i(X)\omega_2),$$

其中 ω_1 为 k 阶微分形式.

b) 证明公式

$$i(X)d + di(X) = L_X,$$

其中 L_X 是沿向量场 X 的李导数.

§26. 反称张量和积分理论

1. 微分形式的积分

如果在 n 维空间的区域 U 上给出了函数 $f(z^1,\cdots,z^n)$, 则在区域上定义了函数 f 的积分:

$$\int\cdots\int_U f(z)dz^1\cdots dz^n=\int\cdots\int_U f(z)dz^1\wedge\cdots\wedge dz^n. \tag{1}$$

这时, 如果给出了坐标变换

$$z=z(y), \tag{2}$$

则成立变量变换公式

$$\int\cdots\int_U f(z)dz^1\wedge\cdots\wedge dz^n=\int\cdots\int_V Jf(z(y))dy^1\wedge\cdots\wedge dy^n, \tag{3}$$

其中 $J=\det\left(\dfrac{\partial z^i}{\partial y^j}\right)$ 为雅可比, 而 V 为原来的区域 U, 但是是在坐标 $y^1,\cdots,y^n$ 下描述的.

在这一节中我们只考虑具有正雅可比的变量变换.

我们注意到, 积分号下的表达式是个 n 阶的反称张量. 在坐标系 $z^1,\cdots,z^n$ 下, 这个张量的分量 $T_{1\cdots n}$ 按定义等于 $f(z)=T_{1\cdots n}$. 回忆定理 18.1 知道, 在 n 维空间中的 n 阶反称张量在坐标变换 $z=z(y)$ 下, 其变换公式为 $T'_{1\cdots n}=J\cdot T_{1\cdots n}$, 其中 $J=\det\left(\dfrac{\partial z^i}{\partial y^j}\right)$. 故而 $T'_{1\cdots n}=Jf(z)$, 这与公式 (3) 吻合. 因此被积表达式是个反称张量.

例 26.1. 如果给出了 (伪) 黎曼度量 (g_{ij}), 则行列式 $g=\det(g_{ij})$ 在变换 $z=z(y)$ 下变成

$$g'=\det(g'_{ij})=\det\left(g_{kl}\frac{\partial z^k}{\partial y^i}\frac{\partial z^l}{\partial y^j}\right)=J^2g. \tag{4}$$

因此表达式 $\sqrt{|g|}$ 在具正雅可比的坐标变换下其表现同于反称张量. 回忆在曲面上区域的面积被定义为 (§7.4):

$$\sigma(U)=\iint_U\sqrt{|g|}dudv,\quad u=z^1,\quad v=z^2,\quad n=2; \tag{5}$$

我们现在看出右端的积分不依赖于坐标的选取.

如果在具欧氏坐标 (x^1,x^2,x^3) 的空间中给出了曲面 $x^i=x^i(z), i=1,2,3$, 并且如果我们要在曲面上对标量函数 $f(x(z))$ 进行某种积分, 而这个函数又与此曲面有本质性的关联 (譬如, 它的高斯曲率), 则这个积分可以这样定义: 在曲面的区域 U 上的积分为

$$\iint_U f(x(z))\sqrt{|g|}dz^1\wedge dz^2, \tag{6}$$

其中 $\iint\limits_U \varphi(z)dz^1 \wedge dz^2$ 为通常的二重积分. 有时称表达式 $\sqrt{|g|}dz^1 \wedge dz^2$ 为曲面上的测度 (面积元).

因此我们得出结论:

1) 在 n 维空间中, 对任意有界区域 U 定义了积分 $\int \cdots \int\limits_U T$, 其中 T 为 $(0,n)$ 型反称张量, $T = (T_{i_1 \cdots i_n})$.

2) 在坐标表示下, 这个张量被表示为

$$T = T_{1\cdots n}dz^1 \wedge \cdots \wedge dz^n$$

(或 $T_{1\cdots n}dz^1 \cdots dz^n$, 这时不写入记号 $\wedge$).

3) $T_{1\cdots n} = f(z)$ 为点的标量函数, 并在坐标变换 $z = z(y)$ 下有

$$\int \cdots \int\limits_U f(z)dz^1 \wedge \cdots \wedge dz^n = \int \cdots \int\limits_V Jf(z(y))dy^1 \wedge \cdots \wedge dy^n,$$

其中 $J = \det\left(\dfrac{\partial z^i}{\partial y^j}\right)$.

4) 如果我们想在空间上对函数 $\varphi(z)$ 积分, 则必须要有事先给出的特定的, 空间中的反称张量 T (称之为体积元或测度). 于是按定义, 函数 $\varphi(z)$ 的积分为张量 $\varphi(z)T$ 的积分

$$\int \cdots \int\limits_U \varphi(z)T = \int \cdots \int\limits_U \varphi(z)T_{1\cdots n}dz^1 \wedge \cdots \wedge dz^n. \tag{7}$$

5) 在此情形下, 如果度量 (g_{ij}) 已给定, 则那个特定反称张量 $T = T_{1\cdots n}dz^1 \wedge \cdots \wedge dz^n$(关于具正雅可比的 "张量") 由形式

$$T = d\sigma = \sqrt{|g|}dz^1 \wedge \cdots \wedge dz^n$$

给出. 这时, 函数 $\varphi(z)$ 的积分被定义为

$$\int \cdots \int\limits_U \varphi(z)T = \int \cdots \int\limits_U \varphi(z)\sqrt{|g|}dz^1 \wedge \cdots \wedge dz^n.$$

6) 记号 $\wedge$ 表明 $dz^i \wedge dz^j = -dz^j \wedge dz^i$. 在变量变换 $z^i = z^i(y)$ 下, 由等式 $dz^j = \dfrac{\partial z^j}{\partial y^i}dy^i, dy^i \wedge dy^j = -dy^j \wedge dy^i$, 我们得到

$$dz^{i_1} \wedge \cdots \wedge dz^{i_k} = \sum_{j_1 < \cdots < j_k} J^{i_1 \cdots i_k}_{j_1 \cdots j_k} dy^{j_1} \wedge \cdots \wedge dy^{j_k}, \tag{8}$$

其中 $J^{(i)}_{(j)}$ 为雅可比矩阵的对应子式. 特别, 当 $k = n$, 我们有

$$dz^1 \wedge \cdots \wedge dz^n = Jdy^1 \wedge \cdots \wedge dy^n, \tag{9}$$

J 为雅可比.

7) 在欧几里得坐标下我们有 $\sqrt{|g|}=1$, 从而

$$d\sigma=dx^1\wedge\cdots\wedge dx^n.$$

需要区别两种表达式.

a) "在区域上对 n 阶反称张量的积分"(在 1)- 3) 中定义), 这是通常的函数的重积分. 它不依赖于 n 维欧氏空间上的度量, 并总是有意义的 ("第 Ⅱ类积分");

b) "在区域上函数的 Ⅰ类积分"(见 4),5)): 要计算这个积分需要知道, 按什么体积元 (测度) 进行积分; 需要将函数乘以这个体积元 (反称张量) 而且只在此后才进行积分. 显然, 这时积分化为了第 Ⅱ类积分.

现在转向在 n 维空间中任意的 k 阶反称张量. 先考察在 n 维空间中的 (0,1)型张量 T_j(余向量). 在 §18 中我们对每个余向量 T_j 构造了对应的一阶微分形式 $\omega=T_jdz^j$.

"微分形式" 这个词与什么有关联? 原来, 表达式 T_jdz^j 可以沿任意曲线 $z^i=z^i(t), a\leqslant t\leqslant b$ 进行积分. 事实上, 考虑表达式

$$\int_a^b T_j\dot{z}^jdt=\int_a^b T_j\xi^jdt,\tag{10}$$

其中 $\xi^j=\dot{z}^j=\dfrac{dz^j}{dt}$ 为速度向量.

称这个表达式为微分形式在曲线上的积分 (在分析中称之为 "第二类积分").

另外的事实是: 给出曲线 $z^i=z^i(t)$ 和某个标量函数 $f(z(t))$ 时, 它与此曲线有实质性的联系. (例如, 它的曲率或挠率). 那么, 引进曲线上的测度, 即长度元 $dl=|\dot{z}|dt$, 从而称表达式 $\displaystyle\int_a^b f(z(t))dt$ 为函数 $f(z(t))$ 沿曲线的积分 (在分析中, 这是 "第一类积分"). 从本质上说, 曲线上的长度元 $dl=|\dot{z}|dt$ 是以前所引进的一般的 "体积元"$d\sigma=\sqrt{|g|}dz^1\wedge\cdots\wedge dz^n$ 的一维情形, 这是因为对 $n=1, |g|=|g_{11}|$, 且 $\sqrt{|g_{11}|}dt=dl$, 其中 $g_{11}=|\dot{z}|^2, t=z^1$.

关于第二类积分, 即沿任意曲线的余向量场的积分, 它具有下面的性质:

1) 它与在曲线上所引进的参数无关:

$$\int_a^b T_\alpha\frac{dz^\alpha}{dt}dt=\int_{a'}^{b'}T_\alpha\frac{dz^\alpha}{d\tau}d\tau,\tag{11}$$

其中 $t=t(\tau)$, 且当 t 由 a 变到 b 时 τ 由 a' 变到 b';

2) 积分的结果也不依赖于空间的坐标: 如果 $z=z(y)$ 和 $T'_\beta=T_\alpha\dfrac{\partial z^\alpha}{\partial y^\beta}$, 又 $z^\alpha(t)=z^\alpha(y(t))$, 则有等式

$$\int_a^b T_\alpha\frac{dz^\alpha}{dt}dt=\int_a^b T'_\beta\frac{dy^\beta}{dt}dt.\tag{12}$$

事实上,$T_\alpha dz^\alpha \equiv T'_\beta dy^\beta$, 而且这两个积分都沿着同一条曲线 $z^\alpha = z^\alpha(t)$ 或 $y^\beta = y^\beta(t)$ 进行, 其中 $z(t) = z(y(t))$; 因此它们相同.

于是, 我们已定义了在 n 维空间中的 (n 维) 区域上 n 阶反称张量的积分和在任意曲线上余向量场的积分.

原来, 这是 $(0,k)$ 型的 k 阶反称张量在 n 维空间中的一个 k 维曲面上的积分. 设 k 维曲面由参数给出,

$$x^i = x^i(z^1, \cdots, z^k), \quad i = 1, \cdots, n. \tag{13}$$

并设有坐标 $z^1, \cdots, z^k$ 的 k 维空间中给出了区域 U. 如何给出 n 维空间中的反称张量 $T = (T_{i_1\cdots i_k})$ 在 k 维空间中区域 U 上的积分? 其中 n 维空间的坐标为 $x^1, \cdots, x^n$, k 维空间中的坐标为 $z^1, \cdots, z^k$, 而后者以 $x^i = x^i(z^1, \cdots, z^k)$ 嵌入在 n 维空间中. 为方便起见, 我们将对反称张量使用微分形式的语言来描述.

在 §22.1 中曾定义过张量 $T_{i_1\cdots i_k}$ 在曲面 $x = x(z)$ 的限制. 它已是 k 维空间 (在曲面上的) 中的 k 阶张量了.

定义 26.1. 设 $(T_{i_1\cdots i_k}) = T$ 为 n 维空间中的张量,U 为曲面 $x^i = x^i(z^1, \cdots, z^k)$ 上的区域. 称 T 在曲面的限制在此区域上通常的重积分为 T 在这个曲面区域 U 上的积分. 其中 $i = 1, \cdots, n$. 它的形式是

$$\int_U \cdots \int T = \int_U \cdots \int \Big(\sum_{i_1 < \cdots < i_k} T_{i_1\cdots i_k} J^{i_1\cdots i_k}_{1\cdots k} \Big) dz^1 \wedge \cdots \wedge dz^k. \tag{14}$$

我们记得, 表达式 $\sum\limits_{i_1 < \cdots < i_k} T_{i_1\cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ 被称做 k 阶微分形式; 我们也已知道这个简明的形式表达了 $(0,k)$ 型的反称张量. 这个对形式 (张量) 在曲面上区域的积分具有两个性质:

1) 在曲面上的变量变换 $z^q = z^q(z'), q = 1, \cdots, k$ 下这个积分不变. 事实上, 限制

$$\left(\sum_{i_1 < \cdots < i_k} T_{i_1\cdots i_k} J^{i_1\cdots i_k}_{1\cdots k} \right) dz^1 \wedge \cdots \wedge dz^k$$

为 k 维空间 $z^1, \cdots, z^k$ 中的 k 阶张量; 在变量变换 $z^q = z^q(z'^1, \cdots, z'^k)$ 下, 它进行了在 k 维空间中区域 U 上的重积分的通常的变量变换.

2) 在这个 n 维空间中的坐标变换 $x^i = x^i(x'^1, \cdots, x'^n)$ 下, 这个积分不变. 就像 $k = 1$ 的情形一样, 由此可立即得到, 在变换 $x = x(x')$ 下有恒等式

$$\sum_{i_1 < \cdots < i_k} T_{i_1\cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k} \equiv \sum_{j_1 < \cdots < j_k} T'_{j_1\cdots j_k} dx'^{j_1} \wedge \cdots \wedge dx'^{j_k}, \tag{15}$$

其中 $dx^i = \dfrac{\partial x^i}{\partial x'^j} dx'^j$, 而分量 $T'_{j_1\cdots j_k}$ 由 $T_{j_1\cdots j_k}$ 按通常的张量规则得到.

在欧氏空间中, 定义曲面 $x^i = x^i(z^1, \cdots, z^k), i = 1, \cdots, n$ 上的黎曼度量: 如果 $x^1, \cdots, x^n$ 为欧氏坐标, 则 $g_{ij}dz^i dz^j = \sum_{q=1}^{n}(dx^q)^2$, 其中 $dx^q = \frac{\partial x^q}{\partial z^\alpha}dz^\alpha$ (参看 §7); 此时,$g_{ij} = g_{ji}, dz^i dz^j = dz^j dz^i$.

像经常做的那样, 曲面上的体积元有形式

$$d\sigma = \sqrt{|g|}dz^1 \wedge \cdots \wedge dz^k; \quad g = \det(g_{ij}). \tag{16}$$

设在曲面上给出一个任意的函数 $f(z^1, \cdots, z^k)$.

定义 26.2. 函数 $f(z^1, \cdots, z^k)$ 在曲面上的 (第 I 类) 积分定义为这个函数对曲面上体积元的积分:

$$\text{第 I 类积分} = \underset{\substack{U \\ (\text{在曲面上})}}{\int \cdots \int} f(z^1, \cdots, z^k)\sqrt{|g|}dz^1 \wedge \cdots \wedge dz^k.$$

需要重视的是, 第 II 类积分不依赖于空间和曲面上的黎曼几何, 而第 I 类积分则通过体积元 $\sqrt{|g|}dz^1 \wedge \cdots \wedge dz^k$ 与它有关联; 而这个体积元是个 k 阶反称张量 (相对于具正雅可比的变换), 它只在所给曲面上由其黎曼度量决定, 而这个曲面上的黎曼度量则由整个空间上的欧氏度量定义.

2. 微分形式的例题

例 26.2. 函数 $f(x)$ 在点 P 的值按定义是 "积分" 的平凡例子, 这是 0 阶张量 (即标量 $f(z)$) 在零维曲面 (点 P) 上的积分:"积分" $= f(P)$. 这个平凡的解释是有所目的的, 它在后面对广义斯托克斯公式的讨论中将被用上.

例 26.3. 余向量场或 1 阶微分形式 $(T_\alpha) = T_\alpha dx^\alpha$ 沿曲线 $x^i = x^i(t), a \leqslant t \leqslant b$ 的积分, 已在前面小节中讨论过.

$$\text{曲线上的 (第 II 类) 积分} = \int_a^b T_\alpha \frac{dx^\alpha}{dt}dt. \tag{17}$$

例 26.4. 张量场 $(T_{ij}) = \sum_{i<j} T_{ij}dx^i \wedge dx^j$ (2 阶微分形式) 在曲面 $x^i = x^i(z^1, z^2)$, $i = 1, \cdots, n$ 的积分有形式

$$\text{在曲面上的积分} = \underset{\substack{U \\ (\text{在曲面上})}}{\iint} \sum_{i<j} T_{ij}dx^i \wedge dx^j, \tag{18}$$

其中 $T_{ij} = T_{ij}(x(z)), dx^i = \frac{\partial x^i}{\partial z^j}dz^j, \quad dz^i \wedge dz^j = -dz^j \wedge dz^i$. 因此,

$$\iint\limits_U \sum_{i<j} T_{ij} dx^i \wedge dx^j = \iint\limits_U \sum_{i<j} T_{ij}(x(z)) \left(\frac{\partial x^i}{\partial z^p} dz^p\right) \wedge \left(\frac{\partial x^j}{\partial z^q} dz^q\right)$$

$$= \iint\limits_U \left[\sum_{i<j} T_{ij} \left(\frac{\partial x^i}{\partial z^1}\frac{\partial x^j}{\partial z^2} - \frac{\partial x^j}{\partial z^1}\frac{\partial x^i}{\partial z^2}\right)\right] dz^1 \wedge dz^2.$$

我们注意到 $\dfrac{\partial x^i}{\partial z^1}\dfrac{\partial x^j}{\partial z^2} - \dfrac{\partial x^j}{\partial z^1}\dfrac{\partial x^i}{\partial z^2} = J_{12}^{ij}$ 为雅可比 $\left(\dfrac{\partial x^i}{\partial z^q}\right)$ 的子式,$i = 1, \cdots, n; q = 1, 2$. 因此最后得到

$$\iint\limits_U \sum_{i<j} T_{ij} dx^i \wedge dx^j = \iint\limits_U \left[\sum_{i<j} T_{ij} J_{12}^{ij}\right] dz^1 \wedge dz^2. \tag{19}$$

在三维空间 $(n = 3)$ 和欧氏坐标 x^1, x^2, x^3 中, 有 $(dl)^2 = \sum (dx^i)^2$, 这个积分通常可表示为

1) 余向量场在曲线上的积分

$$\int_a^b T^\alpha \frac{dx^\alpha}{dt} dt = \int_P^Q T\xi dt, \tag{20}$$

其中 $\xi = \dot{x}, T = (T_\alpha) = (T^\alpha)$ (在欧氏坐标下向量和余向量是相合的概念, 它相对于旋转是不变的), $P = (x^1(a), x^2(a), x^3(a)), Q = (x^1(b), x^2(b), x^3(b)), a \leqslant t \leqslant b, T\xi = \langle T, \xi\rangle$ 为内积;

2) (0,2) 型反称张量 (或 2 阶形式) 在曲面 ("流") 上的积分. 在欧氏坐标 x^1, x^2, x^3 的三维欧氏空间中,$x^i = x^i(z^1, z^2), i = 1, 2, 3$ 定义了一个曲面, 而在此曲面的每点上定义了基本切向量:

$$\xi = (\xi^i) = \left(\frac{\partial x^i}{\partial z^1}\right) = \frac{\partial x^i}{\partial z^1} e_i,$$

$$\eta = (\eta^i) = \left(\frac{\partial x^i}{\partial z^2}\right) = \frac{\partial x^i}{\partial z^2} e_i;$$

它们的向量积 $[\xi, \eta]$ 为曲面的法方向. 向量积实质上是张量 $T^{ij} = (\xi^i \eta^j - \xi^j \eta^i)$, 它相伴于 "向量" (T^i):

$$T^1 = T^{23}, \quad T^2 = -T^{13}, \quad T^3 = T^{12}; \quad T = (T^1, T^2, T^3).$$

显然, 此时有 $T^{ij} = J_{12}^{ij}$.

向量 $[\xi, \eta]$ 指向曲面的法方向. 它的长度为 $\sqrt{g}$, 其中 $g = \det(g_{ij})$,

$$g_{ij} dz^i dz^j = \sum_{i=1}^3 (dx^i)^2; \quad dx^i = \frac{\partial x^i}{\partial z^j} dz^j$$

(见 §7). 因此在具欧氏坐标 (x^1,x^2,x^3) 的欧氏空间中曲面 $x^i=x^i(z^1,z^2)$ 的区域 U 上的积分以下面形式进行:

$$\begin{aligned}\iint\limits_U \sum_{i<j} T_{ij}dx^i\wedge dx^j &= \iint\limits_U \Big(\sum_{i<j} T_{ij}J^{ij}_{12}\Big)dz^1\wedge dz^2\\ &= \iint\limits_U \langle T,[\xi,\eta]\rangle dz^1\wedge dz^2 = \iint\limits_U \langle T,n\rangle\sqrt{|g|}dz^1\wedge dz^2,\end{aligned}$$

其中 n 为单位法向量:

$$n=\frac{[\xi,\eta]}{|[\xi,\eta]|}=\frac{[\xi,\eta]}{\sqrt{|g|}}.$$

注 在 $n=4$ 的四维情形, 对曲面上的 2 阶形式的积分 $(k=2)$ 不能化为对向量上的运算, 即便是欧氏空间也不行.

关于三维情形我们已证明了

定理 26.1. 在三维欧氏空间中对 2 阶形式在曲面上的积分等同于下面的第 I 类型积分:

$$\iint\limits_U \sum_{i<j} T_{ij}dx^i\wedge dx^j = \iint\limits_U \langle T,n\rangle\sqrt{|g|}dz^1\wedge dz^2, \tag{21}$$

其中 U 为曲面上的区域, 在欧氏坐标 (x_1,x_2,x_3) 下这个曲面为 $x^i=x^i(z^1,z^2)$, $i=1,2,3$ 所定义, 而 n 为曲面的单位法向量, T 为在欧氏坐标 (x^1,x^2,x^3) 下张量 (T_{ij}) 所相伴的向量, $g_{ij}dz^idz^j=\sum\limits_{i=1}^{3}(dx^i)^2, dx^i=\dfrac{\partial x^i}{\partial z^j}dz^j$.

在三维的欧氏坐标 (x^1,x^2,x^3) 下由于 $(0,2)$ 型反称张量(形式)和向量之间的关联, 也由于向量和余向量之间的关联, 则可以谈论对向量场 $(T_\alpha)=(T^\alpha)$ 的下列积分:

a) 在闭曲线 Γ 上的: $\oint\limits_\Gamma T_\alpha dx^\alpha$;

b) 在闭曲面 U 上的: $\iint\limits_U \langle T,n\rangle\sqrt{|g|}dz^1\wedge dz^2$, 其中 n 为曲面的单位法向量.

回忆在分析中的两个定义:

1) 如果曲线 Γ 为闭 (即 Γ 具形式 $x^i(t), a\leqslant t\leqslant b$, 且 $x^i(a)=x^i(b), i=1,2,3$), 则对余向量场的积分

$$\oint\limits_\Gamma T_\alpha \dot{x}^\alpha dt$$

被称为沿曲线 Γ 的旋转场.

2) 如果曲面 $U=\{f(x^1,x^2,x^3)=0\}$ 为闭的意义是说它为区域 $f(x^1,x^2,x^3)\leqslant 0$ 的边缘，而此区域在空间中有界，则称积分

$$\iint\limits_U \sum_{i<j} T_{ij}dx^i \wedge dx^j$$

为向量场 $(T_{ij})=-(T_{ji})$ 通过此曲面的总流量，或者在欧几里得情形下，是通过此曲面的向量场 $T=(T^1,T^2,T^3)$ 的流量;$T^1=T_{23},T^2=-T_{13},T^3=T_{12}$; 而 x^1,x^2,x^3 为欧氏坐标.

如果曲面由方程 $f(x^1,x^2,x^3)=0$ 给出，则可能此曲面不能在整体上由 $x^i=x^i(z^1,z^2)$ 这样的参数方程给出. 但是在每个非异点的邻近这件事却可以做到 (见 §7.1). 这个积分不依赖于曲面上坐标的选取. 因此在计算在整个曲面上的积分时，需要将它分割成一些片段，使得每一片都可单独地由参数表示; 然后再取在这些片上积分的和. 例如，球面 $(x^1)^2+(x^2)^2+(x^3)^2=1$ 可以分成两个这样的片: 上下两个半球.

例 26.5. 设在 n 维欧氏空间中给出了超曲面 M^{n-1}:

$$F(x^1,\cdots,x^n)=0,\quad \text{grad } F\neq 0, \tag{22}$$

或者局部地 $x^\alpha=x^\alpha(y^1,\cdots,y^{n-1})$. 这里的 $x^1,\cdots,x^n$ 为欧氏坐标. 定义“曲率形式”

$$Kd\sigma=K\sqrt{|g|}dy^1\wedge\cdots\wedge dy^{n-1}, \tag{23}$$

其中 K 为曲率 (对 $n=2$ 为平面中曲线的曲率，对 $n=3$，这是高斯曲率，对于 $n>3$ 我们没有给出定义，这是由于这种情形对我们并不重要).

考虑球面 S^{n-1}，它由 $\sum(x^\alpha)^2=1$ 给出. 在球面上定义有旋转不变的 $(n-1)$ 维体积元 Ω_{n-1}，在 $n=2,3$ 时分别具形状:

$$n=2,\quad \Omega_{n-1}=d\varphi,$$

$$n=3,\quad \Omega_{n-1}=\sin\ \theta d\theta d\varphi. \tag{24}$$

我们定义由曲面 M^{n-1} 到球面 S^{n-1} 的高斯球面映射如下: 考虑在曲面 M^{n-1} 的点 P 上曲面唯一的法线 n_P，并将此向量 n_P 转移至原点.n_P 的终点则是球面 S^{n-1} 上的点. 对应 $P\mapsto n_P$ 定义了高斯映射:

$$\psi: M^{n-1}\to S^{n-1} \tag{25}$$

(点转变到向量 n_P 的终点，而 n_P 由坐标原点出发).

定理 26.2. 成立下面公式:

$$Kd\sigma=\psi^*(\Omega_{n-1}),$$

特别, $\begin{cases} Kdt = \psi^*(d\varphi), & \text{当} n = 2, \\ K\sqrt{|g|}dy^1 \wedge dy^2 = \psi^*(\Omega)\ (\Omega = \sin\ \theta d\theta d\varphi), & \text{当}\ n = 3; \end{cases}$

其中 $d\sigma = \sqrt{|g|}dy^1 \wedge \cdots \wedge dy^{n-1}$ 为曲面上在局部坐标 $y^1, \cdots, y^{n-1}$ 下的 $(n-1)$ 维体积元.

(运算 ψ^* 在 $(0,k)$ 型张量上的定义见 §22.)

证明 (仅对 $n\geqslant 2$ 的情形). 我们对 $n = 3$ 的情形进行证明. 在 $\mathbb{R}^3$ 中我们选取以 P 为原点的欧氏坐标, 使得 $z = x^3$ 轴在 P 点正交于此曲面, 而轴 $x = x^1, y = x^2$ 切于此曲面. 于是 $y^1 = x, y^2 = y$, 从而在 P 点附近曲面由方程 $z = f(x,y)$ 给出, 它满足 $df|_P = 0$. 又, $K = \det\begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy}\end{pmatrix}, g_{ij} = \delta_{ij}$ 于点 P 成立 (这时 $f_x = f_y = 0$). 在球面 $S^{n-1} = S^2$ 上我们选取那样的坐标, 使得在点 $\psi(P) = Q$ 处 $\bar{z}$ 轴垂直于 S^2, 而 $\bar{x}$ 轴和 $\bar{y}$ 轴切于 S^2; 那么形式 $\Omega = \sin\ \theta d\theta d\varphi$ 在点 Q 有形式 $\Omega = d\bar{x} \wedge d\bar{y}$, 而在点 Q 附近球面由方程 $\bar{z} = \sqrt{1 - \bar{x}^2 - \bar{y}^2}$; 在点 Q 的球面度量有形 $g_{ij} = \delta_{ij}$.

在点 P 附近的点 P' 的法向量的坐标为

$$\left(\frac{-f_x}{\sqrt{1+f_x^2+f_y^2}}, \frac{-f_y}{\sqrt{1+f_x^2+f_y^2}}, \frac{1}{\sqrt{1+f_x^2+f_y^2}}\right)$$

(又, 在点 P 有 $f_x = f_y = 0$). 因此在点 P 附近, 高斯映射 ψ 可记为

$$\bar{x} = \frac{-f_x}{\sqrt{1+f_x^2+f_y^2}}, \quad \bar{y} = \frac{-f_y}{\sqrt{1+f_x^2+f_y^2}} \tag{26}$$

(点 P' 的坐标 x, y 变为点 Q' 的坐标 $\bar{x}, \bar{y}$).

由定义, 形式 $\psi^*(\Omega)$ (参看 §22) 为

$$\begin{aligned}\psi^*(\Omega)|_P &= \left(\frac{\partial\bar{x}}{\partial x}\frac{\partial\bar{y}}{\partial y} - \frac{\partial\bar{x}}{\partial y}\frac{\partial\bar{y}}{\partial x}\right)\bigg|_P dx \wedge dy \\ &= Jdx \wedge dy,\end{aligned} \tag{27}$$

其中 J 为映射 ψ 在点 P 的雅可比. 因为在点 P 有 $f_x = f_y = 0$, 则显然 $J = f_{xx}f_{yy} - f_{xy}f_{yx} = K$ (高斯曲率). 因为在点 P 的模 $|g|$ 等于 1(见 §7.3 的公式 (15) 在曲面上诱导的度量的情形), 于是在所选坐标系 (点 P 处) 下最终成立公式

$$K\sqrt{|g|}dx \wedge dy = \psi^*(\Omega_{n-1}), \quad n = 3. \tag{28}$$

在 $n = 3$ 时的定理已证毕. 对其余剩下的 n, 其证明完全类似. □

注 当 $n = 2$ 时有 $\Omega = d\varphi$ 且对曲线 $x^1 = x^1(y), x^2 = x^2(y)$ 形式 $K\sqrt{|g|}dy$ 为 $K\sqrt{|g|}dy = Kdl$, 其中 dl 为长度元 (l 为自然参数).

3. 广义斯托克斯公式. 例题

我们在前面的一节中已经看到, 对 n 维空间中的 k 维曲面上对 k 阶形式的积分定义并不一定要求这个曲面有形如 $x^i = x^i(z^1, \cdots, z^k)$ 的整体的参数表示. 因为对空间中和平面上的坐标变换, 积分不变, 并且在区域的并上的积分等于积分的和, 故而可将曲面分割为许多片段, 然后在每一块上分别地引进坐标. 在此之后, 在每一块片段上进行积分, 并需对这些结果取和, 从而得到在曲面上的积分.

另外的注解是, 在曲面上常常可以引进那样的整体坐标 $z^1, \cdots, z^k$, 它具有奇点 (见第一章和第二章), 且奇点集合的维数较小, 使在此集合上的积分没有给出任何贡献. 这样的坐标常常被用到积分理论之中. 例如, 其中有平面上的极坐标 (奇点为 $r = 0$), 空间中的柱或球面坐标:(r, φ, z) 为柱坐标 (奇点由直线 $r = 0$ 组成),(r, θ, φ) 为球面坐标 (奇点也由一条直线和对应于坐标值为 $r = 0, \theta = 0, \theta = \pi$ 的那些点充满). 在球面上我们有坐标 (θ, φ), 其奇点为 $\theta = 0, \theta = \pi$. 球面, 这个最简单的曲面, 在它上面原则上不能引进一个单一的没有奇点的坐标系. 在这些例子中坐标系奇点的集合相当小, 它对于积分没有影响. 故而我们对它们 "不予理睬".

在分析中已知曲面上的积分和在它的边界上的积分之间有所关联 ($n = 2$ 时的格林定理,$n = 3$ 和 $k = 3$ 时的高斯 - 奥斯特罗格拉茨基公式, 以及 $n = 3, k = 2$ 的斯托克斯公式). 我们如今以反称张量 (微分形式) 的观点来考虑这些公式.

由于积分的可加性只需知道在每个曲面片段上的基本定义即可. 设在 k 维空间中以不等式 $f(z^1, \cdots, z^k) \leqslant 0$ 的方式给出了曲域 U, 并设 Γ 为方程 $f(x^1, \cdots, x^k) = 0$ 定义的它的边界. 又假设这个有边界的区域在 n 维空间 $x^1, \cdots, x^n$ 中的嵌入由

$$x^i = x^i(z^1, \cdots, z^k), \quad i = 1, 2, \cdots, n$$

给出. 我们于是得到参数化的所给曲面, 其区域和它的边界 Γ 为 $k-1$ 维的曲面.

在区域 U 上的和在其边界上的对所有 n 维空间 $(x^1, \cdots, x^n)$ 中具相应阶数的所有可能的微分形式, 其积分之间究竟有什么联系?

最简单的情形即 $k = 1$ 的情形; 这时 $x^i = x^i(t), z^1 = t$ 为曲线, 区域 U 为线段 $a \leqslant t \leqslant b$, 边界 Γ 是一对点 $t = a$ 和 $t = b$, 又 (纯粹形式地) 由带负号的点 a 走向带正号的点 b.

特别要提及平凡情形, 即在 0 维 "曲面" 上对标量 ("0 阶形式") 的 "积分"; 按定义这个 "曲面" 由带符号的几个点组成.

"维数 0 的曲面" 是点的形式和 $\sum \pm P_i$, 其中 P_i 为空间中的点. 函数 $f(x)$ 在 "0 维曲面" 上的 "积分" 由定义是这个函数在这些点的值以相应的符号的和:

$$\text{此积分} = \sum_i \pm f(P_i).$$

如果在空间中给出了曲线 $x^i(t)$ 和其上的线段 U $(a \leqslant t \leqslant b)$, 其边界为 $\Gamma = Q - P$, 由

分析中我们知道有公式 (牛顿 – 莱布尼茨)

$$\int\limits_{\Gamma} f = f(Q) - f(P) = \int\limits_{U} df = \int_a^b \frac{\partial f}{\partial x^i}\frac{dx^i}{dt}dt, \tag{29}$$

$$P = (x^i(a)), \quad Q = (x^i(b)).$$

这是最简单的"斯托克斯公式", 它将在边界上的积分和在该区域的积分联系起来. 在一种普遍认同的意义下, 高维的斯托克斯型的公式是它的直接推广, 并且进而可以形式地化回到这个最简单的情形.

回到一般的情形, 即在曲面 $x^i = x^i(z^1, \cdots, z^k), i = 1, \cdots, n$ 上有坐标 $z^1, \cdots, z^k$ 的区域 U, 并有由 $f(x^1, \cdots, x^n) = 0$ 给出的边界 Γ (区域 U 由不等式 $f(z) \leqslant 0$ 给出).

如果在空间 $x^1, \cdots, x^n$ 中给出了 $k-1$ 阶形式 (即 $(0, k-1)$ 型反称张量), 它可写为 $T = \sum\limits_{i_1 < \cdots < i_{k-1}} T_{i_1 \cdots i_{k-1}} dx^{i_1} \wedge \cdots \wedge dx^{i_{k-1}}$, 于是可以对它在 $(n-1)$ 维曲面 Γ, 即曲面 $x^i = x^i(z^1, \cdots, z^k), i = 1, \cdots, n$ 上区域 U 的边界上进行积分.

成立一般的定理:

定理 26.3. 对任意微分形式

$$T = \sum_{i_1 < \cdots < i_{k-1}} T_{i_1 \cdots i_{k-1}} dx^{i_1} \wedge \cdots \wedge dx^{i_{k-1}},$$

其中系数 $T_{i_1 \cdots i_k}(x)$ 为光滑, 及任意光滑曲面 $x^i(z^1, \cdots, z^k)$ 和它上面的有界区域 U, 并具有光滑的边界 Γ, 同时它由一块片段构成, 那么成立等式

$$\pm \int\limits_{\Gamma} T = \int\limits_{U} dT. \tag{30}$$

这个公式的平凡情形 $k = 1, k - 1 = 0$ 已在上面指出过. 这里的 dT 为 k 阶形式 (张量 T 或 $k-1$ 阶微分形式的反称的梯度).

在二维和三维情形这个对不同 k 的公式被命名为格林,高斯 – 奥斯特罗格拉茨基以及斯托克斯公式; 这些公式 (30) 的特殊情形独立地在分析课程中给予了证明. 我们来考察这些情形.

1) *平面情形*$(n = 2)$. 给出在平面中的闭曲线 $\Gamma : x^i = x^i(t), x^i(a) = x^i(b)$. 设这条曲线围成了平面上区域 U. 对任意的 (余) 向量场 $T_\alpha dx^\alpha$(一阶形式) 定义了在曲线 Γ 上的积分. 如果向量场 $T_\alpha dx^\alpha$ 在区域 U 内部有定义并且没有奇点, 则成立公式

$$\int\limits_{\Gamma} T_\alpha \frac{dx^\alpha}{dt} dt = \iint\limits_{U} \left(\frac{\partial T_2}{\partial x^1} - \frac{\partial T_1}{\partial x^2}\right) dx^1 \wedge dx^2,$$

或者

$$\int\limits_{\Gamma} (T_1 dx + T_2 dy) = \iint\limits_{U} \left(\frac{\partial T_2}{\partial x} - \frac{\partial T_1}{\partial y}\right) dx \wedge dy. \tag{31}$$

这就是格林公式 (广义斯托克斯定理的特殊情形).

下面补充的是关于所考虑平面为复变量 $z = x + iy$ 的平面情形. 设 $f(z) = f(x,y) = u(x,y) + iv(x,y)$ 为复值函数. 考虑积分

$$\oint_{\Gamma} f(z)dz = \oint_{\Gamma}(u+iv)(dx+idy) = \oint_{\Gamma}(udx - vdy) + i\oint_{\Gamma}(vdx+udy).$$

运用格林公式有

$$-\oint_{\Gamma} f(z)dz = \iint_{U}\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)dx\wedge dy + i\iint_{U}\left(\frac{\partial v}{\partial y}-\frac{\partial u}{\partial x}\right)dx\wedge dy.$$

我们得出结论: 如果函数 $f(z)$ 在区域内部处处光滑, 并且满足柯西 - 黎曼条件 (参看 §12)

$$\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x},\quad \frac{\partial v}{\partial y} = \frac{\partial u}{\partial x}. \tag{32}$$

于是形式 $f(z)dz$ 为闭.

由此事实得到: 如果 n 为非负整数, 则 $\oint_{\Gamma} z^n dz = 0$ 对任意闭周线 Γ 成立; 如果 n 为负, 则 $\int_{\Gamma} z^n dz = 0$ 对不包围 0 点的周线成立, 并且 $\oint_{\Gamma} z^n dz$ 不依赖于 Γ, 这时 Γ 绕 0, 譬如, 按逆时针方向进行. 取 Γ 作为单位圆 $z(t) = e^{it}, 0\leqslant t\leqslant 2\pi$, 由简单计算得到

$$\oint_{\Gamma} z^n dz = \begin{cases} 0, & \text{如果 } n\neq -1 \\ 2\pi i, & \text{如果 } n = -1. \end{cases}$$

这基于重要的 "残数定理".就是说, 对于一致收敛的级数 $f(z) = \sum\limits_{n-\infty}^{\infty} c_n(z-a)^n$ 成立有公式 (周线 Γ 绕点 a 并位于级数的一致收敛的区域内)

$$\begin{aligned} &\oint_{\Gamma} f(z)dz = 2\pi i c_{-1}, \\ &\oint_{\Gamma}(z-a)^{-k}f(z)dz = 2\pi i c_{k-1}. \end{aligned} \tag{33}$$

这些公式允许对解析函数 $f(z)$ 展成泰勒级数时的系数 (如果所有的幂次 $n\geqslant 0$) 或者洛朗级数 $(-\infty < n < \infty)$ 的系数通过积分计算出来.

2) 三维情形. 设给出了三维空间, 其坐标为 x^1, x^2, x^3. 这里要分开两种情形: $k = 2$ 和 $k = 3$.

a) 设 $k=3, U$ 为区域,Γ 为其边界. 于是成立公式

$$\iint\limits_{\Gamma}\sum_{i<j}T_{ij}dx^i\wedge dx^j=\iiint\limits_{U}\left(\frac{\partial T_{12}}{\partial x^3}+\frac{\partial T_{23}}{\partial x^1}-\frac{\partial T_{13}}{\partial x^2}\right)dx^1\wedge dx^2\wedge dx^3. \tag{34}$$

如果 x^1,x^2,x^3 为欧氏坐标, 令 $T^1=T_{23},T^2=-T_{13},T^3=T_{12}$, 则可以定义向量 T. 如果进一步设 n 为曲面的单位法向量, 则根据定理 26.1 我们得到

$$\iint\limits_{\Gamma}\sum_{i<j}T_{ij}dx^i\wedge dx^j=\iint\limits_{\Gamma}\langle T,n\rangle\sqrt{|g|}dz^1\wedge dz^2, \tag{35}$$

其中 z^1,z^2 为曲面上的坐标,$d\sigma=\sqrt{|g|}dz^1\wedge dz^2$ 为其上的面积元.

另外,

$$\frac{\partial T_{12}}{\partial x^3}-\frac{\partial T_{13}}{\partial x^2}+\frac{\partial T_{23}}{\partial x^1}=\frac{\partial T^i}{\partial x^i}=\operatorname{div}T.$$

最后得到

$$\iint\limits_{\Gamma}\langle T,n\rangle\sqrt{|g|}dz^1\wedge dz^2=\iint\limits_{\Gamma}\langle T,n\rangle d\sigma=\iiint\limits_{U}(\operatorname{div}T)dx^1\wedge dx^2\wedge dx^3. \tag{36}$$

这是在三维空间中的高斯–奥斯特罗格拉茨基公式.

b) 设 $k=2,U$ 为曲面 $x^i=x^i(z^1,z^2),i=1,2,3$ 上的区域,Γ 为这个区域的边界(曲线). 我们有

$$\oint\limits_{\Gamma}T_\alpha dx^\alpha=\iint\limits_{U}\left[\left(\frac{\partial T_2}{\partial x^1}-\frac{\partial T_1}{\partial x^2}\right)dx^1\wedge dx^2+\left(\frac{\partial T_3}{\partial x^1}-\frac{\partial T_1}{\partial x^3}\right)dx^1\wedge dx^3+\right.$$

$$\left.\left(\frac{\partial T_3}{\partial x^2}-\frac{\partial T_2}{\partial x^3}\right)dx^2\wedge dx^3\right]. \tag{37}$$

在欧氏情形, 不需要区分向量和余向量, 因而可以将反称张量 (T_{ij}) 与向量 $T=(T^i)$ 看作一样, 那么我们得到了 (斯托克斯公式)

$$\oint\limits_{\Gamma}T_\alpha dx^\alpha=\iint\limits_{U}\langle\operatorname{rot}T,n\rangle dz^1\wedge dz^2, \tag{38}$$

其中 $\operatorname{rot}T$ 是与反称张量 $(\operatorname{rot}T)_{\alpha\beta}=\dfrac{\partial T_\beta}{\partial x^\alpha}-\dfrac{\partial T_\alpha}{\partial x^\beta}$ 相对照的向量.

因此, 在所有这些情况中, 借助于定理 26.1, 广义的斯托克斯公式便转化为各种不同的分析教程中的积分公式. 从而对三维空间已证明了它.

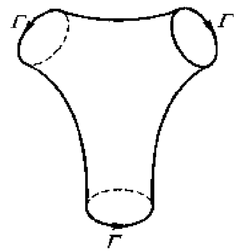


图 31

最后我们注意到, 在广义斯托克斯定理的公式化叙述中没有必要认定边界 Γ 只由一块片段组成. 如果边界 Γ 由几块片段构成 (图 31), 则在不同的块上的积分带有符号 + 或 −, 即需要选取方向. 以前在 $k=1$ 的平凡情形已经注意到了这一点: 那时曲线段的边界 Γ 由两个点组成, 一个有符号 +(终点), 而另一个 (起点) 具符号 −. 在此应注意到, 在表达式 $\iint\limits_{\Gamma}\langle T,n\rangle\sqrt{|g|}dz^1\wedge dz^2$ 中的符号选择 (或者是所说的边界 Γ 的 "定向") 由单位法向量 n 的方向决定 (参看后面的 §26.4).

我们现在指出广义斯托克斯定理的一个应用.

考虑四维空间 $x^0=ct,x^1,x^2,x^3$, 其度量 $g_{ij}=\begin{pmatrix}+1 & & & 0\\ & -1 & & \\ & & -1 & \\ 0 & & & -1\end{pmatrix}$, 其中 c 为光速,t 为时间.

设 $F_{ik}=-F_{ki}$ 为电磁场张量,$i,k=0,1,2,3$. 我们现将研究这个张量在时间 $x^0=ct$ 不变而只允许进行空间坐标 $x^i,i=1,2,3$ 的变换

$$x'^0=x^0,\quad x'^i=x'^i(x^1,x^2,x^3)$$

时的性质. 此时, 张量 (F_{ik}) 在四维空间中定义了电场余向量 $E_\alpha=F_{0\alpha},\alpha=1,2,3,$ 以及磁场张量 $H_{\beta\alpha}=-H_{\alpha\beta}=F_{\alpha\beta},\alpha,\beta=1,2,3$. 如果坐标 x^1,x^2,x^3 是欧几里得的, 则磁场定义了磁场向量 (§21)

$$H^1=H_{23},\quad H^2=-H_{13},\quad H^3=H_{12}.$$

以微分形式表示为

$$d(F_{ij}dx^i\wedge dx^j)=0\quad (\text{麦克斯韦第一对方程})$$

或者在三维欧氏表示中则为

a) $\operatorname{div}\mathbf{H}=\dfrac{\partial H^i}{\partial x^i}=0;$

b) $\operatorname{rot}\mathbf{E}+\dfrac{1}{c}\dfrac{\partial \mathbf{H}}{\partial t}=0.$

由方程 a) 和高斯 – 奥斯特罗格拉茨基公式得到 (其中 Γ 为区域 U 的边界)

$$\iiint\limits_U \operatorname{div} \mathbf{H} dx^1 \wedge dx^2 \wedge dx^3 = \iint\limits_\Gamma \langle \mathbf{H}, \mathbf{n} \rangle d\sigma = 0$$

("穿过闭曲面的磁场流总为零").

由方程 b) 和斯托克斯公式我们得到

$$\iint\limits_U \langle \operatorname{rot} \mathbf{E}, \mathbf{n} \rangle d\sigma = \oint\limits_\Gamma E_\alpha dx^\alpha$$

(Γ 为曲面上区域 U 的边界),

$$-\iint\limits_U \langle \frac{1}{c} \frac{\partial \mathbf{H}}{\partial t}, \mathbf{n} \rangle d\sigma = \oint\limits_\Gamma E_\alpha dx^\alpha$$

("通过曲面的磁场流对时间的导数的负号值等于电场在曲面边界旋转量").

第二对麦克斯韦方程的形状为 (§25)

$$\sum_{k=1}^{3} \frac{\partial F^{ik}}{\partial x^k} + \frac{1}{c} \frac{\partial F^{i0}}{\partial t} = -\frac{4\pi}{c} j^i_{(4)},$$

其中 $j_{(4)} = (c\rho, j_1, j_2, j_3)$, ρ 为电荷密度, $\mathbf{j} = (j_1, j_2, j_3)$ 为三维电流向量 (在度量 $\begin{pmatrix} +1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & -1 \end{pmatrix}$ 下的张量 F_{ik} 的发散量等于 (带符号) 四维的电流向量乘以 $4\pi/c$)

在三维形式中这给出了:

a) $\operatorname{div} \mathbf{E} = 4\pi\rho$,

b) $\operatorname{rot} \mathbf{H} - \dfrac{1}{c} \dfrac{\partial \mathbf{E}}{\partial t} = \dfrac{4\pi}{c} \mathbf{j}$.

由 a) 和高斯 – 奥斯特罗格拉茨基公式我们有

$$\iiint\limits_U 4\pi\rho dx^1 \wedge dx^2 \wedge dx^3 = \iint\limits_\Gamma \langle \mathbf{E}, \mathbf{n} \rangle d\sigma$$

("通过区域边界的电场流等于在整个区域上的总电荷的 4π 倍"). 由 b) 和斯托克斯定理, 我们有

$$\iint\limits_U \langle \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t}, \mathbf{n} \rangle d\sigma + \iint\limits_U \frac{4\pi}{c} \langle \mathbf{j}, \mathbf{n} \rangle d\sigma = \oint\limits_\Gamma H_\alpha dx^\alpha,$$

其中 Γ 为曲面上区域 U 的边界 ("通过曲面的总电流加上通过这个曲面的电场流的导数等于磁场在这个曲面边缘上的旋转量").

我们在这里看到了麦克斯韦第一对和第二对方程的不同几何内涵. 第一对方程与任何的空间度量无关; 至于第二对则没有度量便不能描述. 这些方程的通常形式与欧氏坐标紧密相关.

4. 对立方体上的广义斯托克斯定理的证明.

上面已经指出, 经典的格林和高斯 – 奥斯特罗格拉茨基积分公式是广义斯托克斯公式的特殊情形. 在这一小节中我们要证明积分区域为 k 维立方形时的广义斯托克斯公式.

$\mathbb{R}^n$ 中的*奇异立方体* σ 定义为光滑映射 $\sigma: I^k\to\mathbb{R}^n, n\geqslant k$, 其中 I^k 为 k 维的笛卡儿立方体:

$$I^k=\{(x^1,\cdots,x^k)|0\leqslant x^\alpha\leqslant 1\}.$$

方程 $x^\alpha=0$ 和 $x^\alpha=1$ 分别定义了两块 $(k-1)$ 维边界 I_α^- 和 I_α^+(图 32). 以 ∂I^k 表示立方体 I^k 的边界: $\partial I^k=\bigcup\limits_\alpha(I_\alpha^+\bigcup I_\alpha^-)$.

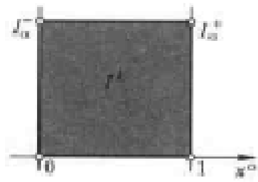


图 32

设 φ^{k-1} 为 $\mathbb{R}^n$ 上 $(k-1)$ 形式,$d\varphi^{k-1}$ 为其外微分; 又设 $\sigma: I^k\to\mathbb{R}^n$ 为奇异立方体.

定理 26.4. 成立等式

$$\int\limits_{\sigma(\partial I^k)}\varphi^{k-1}=\int\limits_{\sigma(I^k)}d\varphi^{k-1}$$

(我们假定边界 ∂I^k 的定向由立方体 I^k 的自然定向诱导. 另外, 诱导是借助于外法线[①]实现的).

注　这里的积分 $\int\limits_{\sigma(\partial I^k)}\varphi$ 是对立方体的全部边界取和.

证明　考虑微分形式 $\omega=\sigma^*(\varphi)$; 由于 σ^* 与运算 d 的可置换性, 我们有 $d\omega=d\sigma^*(\varphi)=\sigma^*(d\varphi)$. 因此对所要求的命题只要证明下面形式即可:

$$\int\limits_{\partial I^k}\omega=\int\limits_{I^k}d\omega.$$

我们在这里应用定义有

$$\int\limits_{\sigma(I^k)}d\varphi=\int\limits_{I^k}\sigma^*(d\varphi)\quad\text{和}\quad\int\limits_{\sigma(\partial I^k)}\varphi=\int\limits_{\partial I^k}\sigma^*(\varphi).$$

①更详细地说, 如果 $n(x)$ 为在点 $x\in\partial I^k$ 处立方体的外法线, 则 ∂I^k 在 x 的诱导定向由标架 $\{v_1,\cdots,v_{k-1}\}$ 定义, 使得 $\{n(x),v_1,\cdots,v_{k-1}\}$ 是 I^k 的已知定向 (因此, 图 32 表出的二维情形中, 所诱导的边界上的定向对应于逆时针旋转方向.)

设 $\omega = a_\alpha(x^1,\cdots,x^k)dx^1\wedge\cdots\wedge\widehat{dx^\alpha}\wedge\cdots\wedge dx^k$, 其中 $a_\alpha(x^1,\cdots,x^k)$ 为光滑函数, 而略去了微分 dx^α. 于是

$$\begin{aligned} d\omega &= \sum_\alpha \frac{\partial a_\alpha}{\partial x^\alpha}dx^\alpha\wedge dx^1\wedge\cdots\wedge\widehat{dx^\alpha}\wedge\cdots\wedge dx^k \\ &= \sum_\alpha(-1)^{\alpha-1}\frac{\partial a_\alpha}{\partial x^\alpha}d^kx, \end{aligned}$$

这里 $d^kx = dx^1\wedge\cdots\wedge dx^k$. 为了简化进一步的讨论, 我们假定函数 $a_\alpha(x^1,\cdots,x^k)$ 表示为乘积形式

$$a_\alpha(x^1,\cdots,x^k) = \prod_{q=1}^k b_\alpha^q(x^q),$$

其中 b_α^q 为单变量 x^q 的光滑函数.(回忆分析教程中曾证明过下面的定理: 任意光滑函数可由单变量光滑函数乘积的线性组合一致逼近; 我们当然不会在这里证明它.)

以显式计算表达式 $\int\limits_{I^k} d\omega$.

$$\begin{aligned} \int\limits_{I^k} d\omega &= \int\limits_{I^k}\left(\sum_\alpha(-1)^{\alpha-1}\frac{\partial a_\alpha}{\partial x^\alpha}\right)d^kx = \int\limits_{I^k}\left(\sum_\alpha(-1)^{\alpha-1}\frac{\partial}{\partial x^\alpha}\left(\prod_{q=1}^k b_\alpha^q(x^q)\right)\right)d^kx \\ &= \int\limits_{I^k}\left(\sum_\alpha(-1)^{\alpha-1}b_\alpha^1(x^1)\cdots\frac{\partial b_\alpha^\alpha(x^\alpha)}{\partial x^\alpha}\cdots b_\alpha^k(x^k)\right)d^kx \\ &= \sum_\alpha\int\limits_{(x^1)}\cdots\int\limits_{(\hat{x}^\alpha)}\cdots\int\limits_{(x^k)}(-1)^{\alpha-1}b_\alpha(x^1)\cdots b_\alpha^{\alpha-1}(x^{\alpha-1})b_\alpha^{\alpha+1}(x^{\alpha+1})\cdots \\ &\quad b_\alpha^k(x^k)\left[(-1)^{\alpha-1}\int\limits_{(x^\alpha)}\frac{\partial b_\alpha^\alpha(x^\alpha)}{\partial x^\alpha}dx^\alpha\right]dx^1\wedge\cdots\wedge\widehat{dx^\alpha}\wedge\cdots\wedge dx^k \\ &= \sum_\alpha\int\limits_{(x^1)}\cdots\int\limits_{(\hat{x}^\alpha)}\cdots\int\limits_{(x^k)}(b_\alpha^1(x^1)\cdots\widehat{b_\alpha^\alpha}(x^\alpha)\cdots b_\alpha^k(x^k)\times[b_\alpha^\alpha(1)-b_\alpha^\alpha(0)]) \\ &\quad dx^1\wedge\cdots\wedge\widehat{dx^\alpha}\wedge\cdots\wedge dx^k \\ &= \sum_\alpha\int\limits_{(x^1)}\cdots\int\limits_{(\hat{x}^\alpha)}\cdots\int\limits_{(x^k)}(b_\alpha^1(x^1)\cdots b_\alpha^\alpha(1)\cdots b_\alpha^k(x^k)- \\ &\quad b_\alpha^1(x^1)\cdots b_\alpha^\alpha(0)\cdots b_\alpha^k(x^k))dx^1\wedge\cdots\wedge\widehat{dx^\alpha}\wedge\cdots\wedge dx^k \\ &= \sum_\alpha\int\limits_{(x^1)}\cdots\int\limits_{(\hat{x}^\alpha)}\cdots\int\limits_{(x^k)}(a_\alpha(x^1,\cdots,x^k)|_{x^\alpha=1}- \\ &\quad a_\alpha(x^1,\cdots,x^k)|_{x^\alpha=0})dx^1\wedge\cdots\wedge\widehat{dx^\alpha}\wedge\cdots\wedge dx^k \\ &= \int\limits_{\partial I^k}\omega. \end{aligned}$$

定理得证. □

习题

26.1. 计算在基灵度量下群 $SO(3,\mathbb{R})$ 和 $SU(2)$ 的体积.

26.2. 设 X^i 为在四维闵可夫斯基空间 $\mathbb{R}^4_1$ 中的向量. 我们定义这个向量场在三维超曲面上的积分 (向量场通过此超曲面的 "流") 为对 3 形式 $X^i dS_i$ 的积分, 其中

$$dS_i = \frac{1}{6}\sqrt{|g|}\varepsilon_{jkli}dx^j \wedge dx^k \wedge dx^l,$$

ε_{jkli} 为 4 阶反称张量 (其定义在 §18.2 中). 试由广义斯托克斯定理推导出等式 (在伪欧氏坐标下)

$$\int\limits_{\partial V} X^i dS_i = \int \frac{\partial X^i}{\partial x^i} dV.$$

26.3. 设 U 为空间中具光滑边界的有界区域, 而空间上给出了黎曼度量 g_{ij}, 又设 Ω^p_U 为所有光滑 p 形式的空间, 这些形式在 U 的外部均为零. 我们在空间 Ω^p_U 中引进内积为

$$\langle \omega_1, \omega_2 \rangle = \int\limits_U \omega_1 \wedge *\omega_2,$$

ω_1, ω_2 为 Ω^p_U 中的 p 形式.

a) 证明空间 Ω^p_U 在所引进的内积为欧几里得的 (参看习题 19.2).

b) 算子 $*$ 为正交:

$$\langle *\omega_1, *\omega_2 \rangle = \langle \omega_1, \omega_2 \rangle.$$

c) 算子 $\delta = (-1)^{np+n+1} * d*$ 与 d 相伴随, 即

$$\langle d\omega_1, \omega_2 \rangle = \langle \omega_1, \delta\omega_2 \rangle.$$

d) δ 的二重算子为零:$\delta\delta = 0$.

e) 设 $\nabla = d\delta + \delta d$. 证明算子 ∇ 在空间 Ω^p_U 中自伴: $\langle \Delta\omega_1, \omega_2 \rangle = \langle \omega_1, \Delta\omega_2 \rangle$. 验证下面的置换关系:

$$\Delta d = d\Delta, \quad \Delta\delta = \delta\Delta, \quad \Delta * = *\Delta.$$

§27. 复空间中的微分形式

1. 算子 d' 和 d''

设 D 为复空间 $\mathbb{C}^n$ 中的区域, 空间的坐标为 $z^1, \cdots, z^n$. 可以考虑 "实化" 区域 $D^{\mathbb{R}}$, 其实坐标为 $x^1, \cdots, x^n, y^1, \cdots, y^n$, 其中

$$z^k = x^k + iy^k, \quad k = 1, \cdots, n. \tag{1}$$

向量 $\frac{\partial}{\partial x^1},\cdots,\frac{\partial}{\partial x^n},\frac{\partial}{\partial y^1},\cdots,\frac{\partial}{\partial y^n}$ 构成 $D^{\mathbb{R}}$ 的切空间的基. 我们将考虑复切空间, 即复系数的线性组合

$$\sum_{k=1}^{n} a_k \frac{\partial}{\partial x^k}+\sum_{k=1}^{n} b^k \frac{\partial}{\partial y^k}. \tag{2}$$

在这些向量构成的空间中引进复基底是方便的, 即基底 $\frac{\partial}{\partial z^k},\frac{\partial}{\partial \bar{z}^k}$, 其中

$$\frac{\partial}{\partial z^k}=\frac{1}{2}\left(\frac{\partial}{\partial x^k}-i\frac{\partial}{\partial y^k}\right),\quad \frac{\partial}{\partial \bar{z}^k}=\frac{1}{2}\left(\frac{\partial}{\partial x^k}+i\frac{\partial}{\partial y^k}\right),k=1,\cdots,n. \tag{3}$$

为了得到进一步更为清晰的公式, 我们给出一些约定. 对应于带“杠”的坐标的张量的指标也常常被画上“杠”, 除非那些记号已经有了“$-$”. 例如, $a_{k\bar{l}}z^k\bar{z}^l$ 是对这两个指标都求和 (参看下面的公式 (4),(7),(9) 等等).

任意复向量 ξ 因而有形状

$$\xi=\xi^k\frac{\partial}{\partial z^k}+\xi^{\bar{k}}\frac{\partial}{\partial \bar{z}^k},\quad \xi=(\xi^k,\xi^k), \tag{4}$$

这里的向量 ξ 为实的当且仅当的情形是 $\xi^{\bar{k}}=\overline{\xi^k}$. 在复余向量 (1 形式) 空间中的对偶基的形状是

$$dz^k=dx^k+idy^k,\quad d\bar{z}^k=dx^k-idy^k,\quad k=1,\cdots,n. \tag{5}$$

任意复 k 形式 $\omega,1\leqslant k\leqslant 2n$ 是表达式 $dz^{i_1}\wedge\cdots\wedge dz^{i_p}\wedge d\bar{z}^{j_1}\wedge\cdots\wedge d\bar{z}^{j_q},p+q=k$ 的线性组合. 因而 ω 可以表示为

$$\omega=\omega_{k,0}+\omega_{k-1,1}+\cdots+\omega_{0,k}, \tag{6}$$

其中 $\omega_{p,q},p+q=k$ 包含了 p 个微分 dz^i 及 q 个微分 $d\bar{z}^j$, 即

$$\omega_{p,q}=\frac{1}{p!q!}\sum_{\substack{i_1\cdots i_p\\ \bar{j}_1\cdots\bar{j}_q}} T_{i_1\cdots i_p\bar{j}_1\cdots\bar{j}_q}dz^{i_1}\wedge\cdots\wedge dz^{i_p}\wedge d\bar{z}^{j_1}\wedge\cdots\wedge d\bar{z}^{j_q}, \tag{7}$$

其中分量 $T_{i_1\cdots i_p\bar{j}_1\cdots\bar{j}_q}$ 对于指标 $i_1,\cdots,i_p$ 和 $j_1,\cdots,\bar{j}_q$ 分别为反称的. 称形式 $\omega_{p,q}$ 为 (p,q) 型形式.

展开式 (6) 不依赖于复坐标的选择. 如果 $w^1,\cdots,w^n$ 为区域中另一组复坐标, 而 $w^i=w^i(z^1,\cdots,z^n)$ 复解析函数, 则 dw^i 只是 $dz^1,\cdots,dz^n$ 的形式的线性组合, 而 $d\bar{w}^i$ 为 $d\bar{z}^1,\cdots,d\bar{z}^n$ 的形式的线性组合.

引理 27.1. 微分 d 可以唯一地表示为

$$d-d'+d'', \tag{8}$$

其中对于 (p,q) 型的形式而言,$d'\omega$ 为 $(p+1,q)$ 型, 而微分 $d''\omega$ 为 $(p,q+1)$ 型. 算子 d',d'' 对于复解析坐标变换不变.

证明 对形状 (7) 的形式 $\omega_{p,q}$ 我们有

$$d\omega_{p,q}=\frac{1}{p!q!}\sum_{\substack{i,i_1,\cdots,i_p\\ j_1,\cdots,j_q}}\frac{\partial T_{i_1\cdots i_p\bar{j}_1\cdots\bar{j}_q}}{\partial z^i}dz^i\wedge dz^{i_1}\wedge\cdots\wedge dz^{i_p}\wedge dz^{j_1}\wedge\cdots\wedge dz^{j_q}+$$

$$\frac{(-1)^p}{p!q!}\sum_{\substack{i_1,\cdots,i_p\\ j,\bar{j}_1,\cdots,\bar{j}_q}}\frac{\partial T_{i_1\cdots i_p\bar{j}_1\cdots\bar{j}_q}}{\partial \bar{z}^j}dz^{i_1}\wedge\cdots\wedge dz^{i_p}\wedge d\bar{z}^j\wedge d\bar{z}^{j_1}\wedge\cdots\wedge d\bar{z}^{j_q}. \tag{9}$$

我们以 $d'\omega_{p,q}$ 表示这个公式中的第一项, 以 $d''\omega_{p,q}$ 表示第二项. 由于展开式 (6) 的唯一性, 在 (p,q) 型形式的和中算子 d' 和 d'' 完全由等式 (9) 所决定. 它的不变性来自展式 (6) 对于复解析变换的不变性. 引理证讫. □

推论 d' 和 d'' 的二重算子为零:

$$(d')^2=(d'')^2=0, \tag{10}$$

又这两个算子间是反交换的:

$$d''d'=-d'd''. \tag{11}$$

证明 由等式 $d^2=0$ 及 $d'+d''=d$ 得到

$$0=d^2\omega_{p,q}=(d')^2\omega_{p,q}+(d'')^2\omega_{p,q}+(d'd''\omega_{p,q}+d''d'\omega_{p,q}).$$

右端第一项为 $(p+2,q)$ 型, 第二项为 $(p,q+2)$ 型, 第三和第四项为 $(p+1,q+1)$ 型. 因此所有这三个形式均为零. 从而得证. □

定义 27.1. 如果

$$d''\omega=0 \tag{12}$$

则称 $(p,0)$ 型形式 ω 为全纯的.

例 27.1. (0,0) 型形式 ω 为复函数 $f(z^1,\cdots,z^n,\bar{z}^1,\cdots,\bar{z}^n)$. 条件 $d''f=0$ 即函数 f 的复解析性的条件:

$$\frac{\partial f}{\partial \bar{z}^i}=0,\quad i=1,\cdots,n.$$

在一般情形下, 如果 $(p,0)$ 型形式的系数为复解析函数, 则形式 ω 为全纯.

2. 凯勒度量. 曲率形式

区域 D 中的埃尔米特度量由函数组 $g_{j\bar{k}}$ 给出, 它参照每个坐标系 $(z^1,\cdots,z^n)$ 满足

a) $g_{j\bar{k}}=\overline{g_{k\bar{j}}}$;

b) 在坐标变换 $z^j=z^j(z^{1'},\cdots,z^{n'})$, $\dfrac{\partial z^j}{\partial \bar{z}^{k'}}\equiv 0$, 我们有

$$g_{j'\bar{k}'}=g_{j\bar{k}}\frac{\partial z^j}{\partial z^{j'}}\overline{\left(\frac{\partial z^k}{\partial z^{k'}}\right)};$$

c) 形式 $g_{j\bar{k}}\xi^j\bar{\xi}^k$ 为正定 (由于 a) 其为实的).

复切向量 $\xi=(\xi^k,\xi^{\bar{k}})$ 和 $\eta=(\eta^k,\eta^k)$ 的复内积由公式

$$\langle\xi,\eta\rangle_{\mathbb{C}}=g_{j\bar{k}}\xi^j\eta^{\bar{k}} \tag{13}$$

给出. 内积 $\langle,\rangle_{\mathbb{C}}$ 不依赖于复坐标的选取. 它在实区域 $D^{\mathbb{R}}$ 中给出了黎曼度量 $\langle,\rangle_{\mathbb{R}}$ 为

$$\langle\xi,\eta\rangle_{\mathbb{R}}=\mathrm{Re}\langle\xi,\eta\rangle_{\mathbb{C}} \tag{14}$$

(参看 §11.2).

对实向量即对 $\xi^{\bar{k}}=\overline{\xi^k},\eta^{\bar{k}}=\overline{\eta^k}$ 的向量 ξ,η, 它满足埃尔米特条件

$$\langle\xi,\eta\rangle_{\mathbb{C}}=\overline{\langle\eta,\xi\rangle_{\mathbb{C}}}. \tag{15}$$

故而表达式

$$\begin{aligned}\{\xi,\eta\}=-\frac{1}{2}\mathrm{Im}\langle\xi,\eta\rangle_{\mathbb{C}}&=\frac{i}{2}\left[\frac{\langle\xi,\eta\rangle_{\mathbb{C}}-\overline{\langle\xi,\eta\rangle}_{\mathbb{C}}}{2}\right]\\&=\frac{i}{2}g_{j\bar{k}}\left(\frac{\xi^j\bar{\eta}^k-\bar{\xi}^k\eta^j}{2}\right)\end{aligned} \tag{16}$$

对 ξ 和 η 是反称的, 从而给出 $D^{\mathbb{R}}$ 上的二阶微分形式. 形式 Ω 的显式表达为

$$\Omega=\frac{i}{2}g_{j\bar{k}}dz^j\wedge d\bar{z}^k; \tag{17}$$

从而 Ω 为 (1,1) 型.

定义 27.2. 如果埃尔米特度量 $g_{j\bar{k}}$ 对应的形式 Ω 为闭, 即

$$d\Omega=0,$$

则称此度量为凯勒度量.

凯勒条件的几何意义将在 §29.4 解释清楚.

例 27.2. 假设在二维实曲面上区域 D 上引进了共形坐标. 在其中的度量有形状

$$ds^2=gdzd\bar{z}. \tag{18}$$

于是 Ω 为曲面的面积元:

$$\Omega=\frac{i}{2}\sqrt{g}(dx+idy)(dx-idy)=\sqrt{g}dx\wedge dy.$$

由于维数的原因,Ω 是闭的, 从而这个度量总是凯勒的.

引理 27.2.　设 $g_{j\bar{k}}$ 为埃尔米特度量,$g=\det(g_{j\bar{k}})$. 令

$$\omega=\frac{1}{2\pi i}d''d'\ln g.$$

那么 ω 为 (1,1) 形式.

证明　只需验证形式 ω 的定义不依赖于复坐标的选取. 作复解析坐标变换

$$z^j=z^j(w^1,\cdots,w^n),\quad \frac{\partial z^j}{\partial \bar{w}^k}=0. \tag{19}$$

设 $J=\det\left(\frac{\partial z^j}{\partial w^k}\right)$. 对应的实坐标变换的雅可比为 $|J|^2$ (见 §12.2). 于是此时埃尔米特形式的矩阵的行列式 $\tilde{g}$ 为

$$\tilde{g}=|J|^2g=J\bar{J}g.$$

由此

$$d'\ln\tilde{g}=d'\ln J+d'\ln\bar{J}+d'\ln g=d'\ln J+d'\ln g,$$

由于函数 J 的解析性, 其中 $d'\ln\bar{J}=0$. 另外,

$$d''d'\ln\tilde{g}=d''d'\ln J+d''d'\ln g=-d'd''\ln J+d''d'\ln g=d''d'\ln g,$$

在其中我们用到了等式 $d'd''=-d''d'$, 而 $d''J=0$. 引理证讫. □

再考虑在共形坐标下的二维实曲面的度量 $gdzd\bar{z}$. 我们有定理:

定理 27.1.　形式 $\omega=\frac{1}{2\pi i}d''d'\ln g$ 为

$$\omega=-\frac{1}{2\pi}K\Omega,$$

其中 K 为曲面的高斯曲率, 而形式 $\Omega=\frac{i}{2}gdz\wedge d\bar{z}$ 为面积元.

证明　我们有

$$d'd''\ln g=\frac{\partial^2}{\partial z\partial\bar{z}}\ln g dz\wedge d\bar{z}.$$

我们在 §13.1 看到, 对于给了共形坐标的曲面而言,$-\frac{2}{g}\frac{\partial^2\ln g}{\partial z\partial\bar{z}}$ 等于高斯曲率. 证完.□

§28. 共变微分

1. 欧氏联络

在 §25 中我们分析过取反称张量 (张量场) 梯度的运算, 它变到阶数增加了 1 的反称张量. 这个运算在分量上为

$$(dT)_{i_1\cdots i_{k+1}}=\sum_{q=1}^{k+1}(-1)^{q+1}\frac{\partial T_{i_1\cdots\hat{i}_q\cdots i_{k+1}}}{\partial x^{i_q}}. \tag{1}$$

特别 $k=1$ 时有

$$(dT)_{ij}=\frac{\partial T_j}{\partial x^i}-\frac{\partial T_i}{\partial x^j}. \tag{2}$$

在那里也曾指出过 dT 仍旧是个张量 (对 $k=0,1$ 有严格的证明). 也指出过算子 d 是唯一的不依赖于几何结构的作用在张量上的微分算子; 这里所说的唯一性的意思是指所有其他这种微分算子都可化为它和它的纯粹的以前所列出的代数运算 (指标置换, 加, 积, 取迹).

关于函数梯度在具笛卡儿坐标 $x^1,\cdots,x^n$ 的空间中的 (p,q) 型张量上按常规的推广

$$T^{i_1\cdots i_p}_{j_1\cdots j_q;k}=\frac{\partial T^{i_1\cdots i_p}_{j_1\cdots j_q}}{\partial x^k}, \tag{3}$$

也已经指出过, 这个运算的结果并不是 $(p,q+1)$ 型的张量. 因为我们会经常遇到这个运算, 我们将指出一类变换, 在这类变换下其变化结果就像张量那样. 这就是线性变换.

定理 28.1. 如果在空间中给出了坐标和张量 $T=T^{i_1\cdots i_p}_{j_1\cdots j_q}$, 则在所有线性坐标变换

$$x^i=a^i_jz^j,\quad i,j=1,\cdots,n,\quad a^i_j=\text{常数},\quad z^i=b^i_jx^j,\quad b^i_ja^j_k=\delta^i_k$$

下, 场

$$T^{i_1\cdots i_p}_{j_1\cdots j_q;k}=\frac{\partial T^{i_1\cdots i_p}_{j_1\cdots j_q}}{\partial x^k}$$

像 $(p,q+1)$ 型张量那样变化.

证明 对线性变换, 我们有

$$\frac{\partial x^i}{\partial z^j}=a^i_j=\text{ 常数},\quad \text{且}\ \frac{\partial^2x^i}{\partial z^j\partial z^k}=0,$$

$$\frac{\partial z^i}{\partial x^j}=b^i_j=\text{ 常数},\quad a^i_jb^j_k=\delta^i_k.$$

按张量的定义

$$\widetilde{T}^{k_1\cdots k_p}_{l_1\cdots l_q} = T^{i_1\cdots i_p}_{j_1\cdots j_q}\frac{\partial z^{k_1}}{\partial x^{i_1}}\cdots\frac{\partial z^{k_p}}{\partial x^{i_p}}\cdot\frac{\partial x^{j_1}}{\partial z^{l_1}}\cdots\frac{\partial x^{j_q}}{\partial z^{l_q}} = T^{(i)}_{(j)}b^{(k)}_{(i)}a^{(j)}_{(l)}, \tag{4}$$

其中 $(i)=i_1\cdots i_p,(k)=k_1\cdots k_p,(j)=j_1\cdots j_q,(l)=l_1\cdots l_q$. 因为 $a^i_j=$ 常数,$b^j_k=$ 常数, 则在微分公式 (4)下, 我们将有

$$\begin{aligned}\widetilde{T}^{(k)}_{(l);r} &= \frac{\partial\widetilde{T}^{(k)}_{(l)}}{\partial z^r} = \frac{\partial T^{(i)}_{(j)}}{\partial z^r}b^{(k)}_{(i)}a^{(j)}_{(l)} = \frac{\partial T^{(i)}_{(j)}}{\partial x^s}\frac{\partial x^s}{\partial z^r}b^{(k)}_{(i)}a^{(j)}_{(l)}\\ &= T^{(i)}_{(j);s}a^s_r a^{(j)}_{(l)}b^{(k)}_{(i)}.\end{aligned} \tag{5}$$

这是张量的变换规则. 定理得证. □

我们在证明中关键地应用了 $\dfrac{\partial^2 x^i}{\partial z^j\partial z^k}=0$. 例如, 考虑 (0,1) 型和 (1,0) 型张量:

$$\frac{\partial T_i}{\partial x^k} = T_{i;k},\quad \frac{\partial T^i}{\partial x^k} = T^i_{;k}.$$

由刚刚证明的定理知,$T_{i;k}$ 和 $T^i_{;k}$ 对于线性坐标变换像张量那样变化. 在一般的坐标变换 $x^i=x^i(z^1,\cdots,z^n),i=1,\cdots,n$, 并设 $\dfrac{\partial^2 x^i}{\partial z^k\partial z^j}\neq 0$ 下, 我们有

$$\begin{aligned}\widetilde{T}_{j;q} &= \frac{\partial\widetilde{T}_j}{\partial z^q} = \frac{\partial}{\partial z^q}\left(T_i\frac{\partial x^i}{\partial z^j}\right) = \frac{\partial T_i}{\partial z^q}\frac{\partial x^i}{\partial z^j} + T_i\frac{\partial^2 x^i}{\partial z^q\partial z^j}\\ &= \frac{\partial T_i}{\partial x^p}\frac{\partial x^p}{\partial z^q}\frac{\partial x^i}{\partial z^j} + T_i\frac{\partial^2 x^i}{\partial z^q\partial z^j} = T_{i;p}\frac{\partial x^p}{\partial z^q}\frac{\partial x^i}{\partial z^j} + T_i\frac{\partial^2 x^i}{\partial z^q\partial z^j}.\end{aligned}$$

这里的 $\widetilde{T}_i$ 是在坐标系 (z) 下的分量, 而 T_i 是在坐标系 (x) 下的分量. 因此, 一般变换公式的形状为

$$\frac{\partial\widetilde{T}_j}{\partial z^q} = \widetilde{T}_{j;q} = T_{i;p}\frac{\partial x^p}{\partial z^q}\frac{\partial x^i}{\partial z^j} + T_i\frac{\partial^2 x^i}{\partial z^q\partial z^j}. \tag{6}$$

$T_i\dfrac{\partial^2 x^i}{\partial z^q\partial z^j}$ 不具有张量的特征. 如同在 §25 证明过的, 表达式

$$\begin{aligned}(d\widetilde{T})_{jq} &= \widetilde{T}_{q;j} - \widetilde{T}_{j;q} = (T_{p;i}-T_{i;p})\frac{\partial x^p}{\partial z^q}\frac{\partial x^i}{\partial z^j}\\ &= (dT)_{ip}\frac{\partial x^p}{\partial z^q}\frac{\partial x^i}{\partial z^j}\end{aligned}$$

为张量. 然而其对称部分

$$(\widetilde{T}_{j;q}+\widetilde{T}_{q;j}) = (T_{i;p}+T_{p;i})\frac{\partial x^p}{\partial z^q}\frac{\partial x^i}{\partial z^j} + 2T_i\frac{\partial^2 x^i}{\partial z^q\partial z^j}$$

不是关于任意坐标变换的张量.

对于张量 T^i 我们有类似的

$$\widetilde{T}^j_{;l}=\frac{\partial \widetilde{T}^j}{\partial z^l}=\frac{\partial}{\partial z^l}\left(T^i\frac{\partial z^j}{\partial x^i}\right)=\frac{\partial T^i}{\partial z^l}\frac{\partial z^j}{\partial x^i}+T^i\frac{\partial}{\partial z^l}\left(\frac{\partial z^j}{\partial x^i}\right)$$

$$=\frac{\partial T^i}{\partial x^p}\frac{\partial x^p}{\partial z^l}\frac{\partial z^j}{\partial x^i}+T^i\frac{\partial^2 z^j}{\partial x^i\partial x^q}\frac{\partial x^q}{\partial z^l}. \tag{7}$$

可清楚看出, 由于第二项这不是一个张量. 由 (7) 得到

$$\widetilde{T}^j_{;j}=\frac{\partial \widetilde{T}^j}{\partial z^j}=\frac{\partial T^i}{\partial x^p}\frac{\partial x^p}{\partial z^l}\frac{\partial z^l}{\partial x^i}+T^i\frac{\partial^2 z^j}{\partial x^i\partial x^q}\frac{\partial x^q}{\partial z^j}$$

$$=T^i_{;p}\delta^p_i+T^i\frac{\partial^2 z^j}{\partial x^i\partial x^q}\frac{\partial x^q}{\partial z^j}=T^i_{;i}+T^i\frac{\partial x^q}{\partial z^j}\frac{\partial^2 z^j}{\partial x^i\partial x^q}. \tag{8}$$

注 $\dfrac{\partial T^i}{\partial x^i}=T^i_{;i}$ (在欧氏坐标下进行计算) 常被称做向量场的发散量. 这个表达式对于非线性坐标变换不是个标量.

这个公式一般只用于欧氏坐标 $(x^1,\cdots,x^n)$, 而发散量的意义在于: 如果给了空间中点的小位移

$$x^i\longmapsto x^i+T^i(x^1,\cdots,x^n)=x^i,$$

则在区域位移后欧氏体积元 $dx^1\wedge\cdots\wedge dx^n$ 得到了增量 $T^i_{;i}dx^1\wedge\cdots\wedge dx^n$ (参看 §22.2).

我们现在返回到梯度

$$T^{i_1\cdots i_p}_{j_1\cdots j_q;k}=\frac{\partial T^{i_1\cdots i_p}_{j_1\cdots j_q}}{\partial x^k} \tag{9}$$

的变换规律. 由于所证的定理, 我们约定只在欧氏坐标 $x^1,\cdots,x^n$ 和在欧氏线性变换

$$x^i=a^i_jx^j,\quad a^i_j=\text{常数}$$

下得到的不同坐标中进行这种运算.

我们已证明, 在另一个不同于 (x) 的非线性变换得到的坐标系中, 按这个公式进行运算时, 我们得到与 $T^{(i)}_{(j);s}$ 相伴的表达式 $\widetilde{T}^{k_1\cdots k_p}_{l_1\cdots l_q;r}$ 是非张量规律的变化.

现在到了以另一种观点看这个问题的时候了: 我们从什么地方知道, 取梯度的运算一定要由同一个公式定义?

可以假定下面的情形:

a) 运算与欧几里得几何应有完全实质性的联系;

b) 只能在欧氏坐标 x 下它才能按公式 (9) 进行;

c) 作为这个运算的结果应得到张量.

由这些假定得出什么推论? 应按什么样的公式这个运算能应用于其他的与欧氏非线性变换有关的坐标系? 对这些假定得出的结论, 我们首先应该计算在欧氏坐标 $x^1,\cdots,x^n$ 下对向量场 T 进行运算的结果, 并且只有在这之后, 根据张量规律, 转换

这个结果到 (任意) 其他坐标系上:$x^i = x^i(z^1, \cdots, z^n), i = 1, \cdots, n$. 利用缩写, 我们令

$$T^{(i)}_{(j);s} = \frac{\partial T^{(i)}_{(j)}}{\partial x^s}, \quad (i) = i_1, \cdots, i_p, \quad (j) = j_1 \cdots j_q. \tag{10}$$

按定义, 我们假定 $T^{(i)}_{(j);s}$ 为张量. 因此有

$$\widetilde{T}^{(k)}_{(l);r} = T^{(i)}_{(j);s} \frac{\partial x^{(j)}}{\partial z^{(l)}} \frac{\partial z^{(k)}}{\partial x^{(i)}} \frac{\partial x^s}{\partial z^r}, \tag{11}$$

其中

$$\frac{\partial x^{(j)}}{\partial z^{(l)}} = \frac{\partial x^{j_1}}{\partial z^{l_1}} \cdots \frac{\partial x^{j_q}}{\partial z^{l_q}}, \quad \frac{\partial z^{(k)}}{\partial x^{(i)}} = \frac{\partial z^{k_1}}{\partial x^{i_1}} \cdots \frac{\partial z^{k_p}}{\partial x^{i_p}}.$$

问题是, 在坐标系 $(z^1, \cdots, z^n)$ 下什么样的运算可以由 $\widetilde{T}^{(k)}_{(l)}$ 得到 $\widetilde{T}^{(k)}_{(l);q}$?

为简明起见考虑向量场 (T^i) 和余向量场 (T_j). 在这种情形下, 由 (11) 得到

$$\widetilde{T}^k_{;q} = T^i_{;s} \frac{\partial z^k}{\partial x^i} \frac{\partial x^s}{\partial z^q}; \quad \widetilde{T}_{l;r} = T_{j;s} \frac{\partial x^j}{\partial z^l} \frac{\partial x^s}{\partial z^r}. \tag{12}$$

因为 $T^i_{;s} = \dfrac{\partial T^i}{\partial x^s}$, 故由公式 (12) 得到

$$\widetilde{T}^k_{;q} = \frac{\partial T^i}{\partial x^s} \frac{\partial x^s}{\partial z^q} \frac{\partial z^k}{\partial x^i} = \frac{\partial T^i}{\partial z^q} \frac{\partial z^k}{\partial x^i}. \tag{13}$$

回想有 $\widetilde{T}^k = T^i \dfrac{\partial z^k}{\partial x^i}$. 由 (13) 得到等式

$$\widetilde{T}^k_{;q} = \frac{\partial T^i}{\partial z^q} \frac{\partial z^k}{\partial x^i} = \frac{\partial}{\partial z^q}(\widetilde{T}^k) - T^i \frac{\partial}{\partial z^q} \left(\frac{\partial z^k}{\partial x^i} \right). \tag{14}$$

由于 $T^i = \widetilde{T}^s \dfrac{\partial x^i}{\partial z^s}$, 于是得到下面的等式

$$\widetilde{T}^k_{;q} = \frac{\partial \widetilde{T}^k}{\partial z^q} - \widetilde{T}^s \frac{\partial x^i}{\partial z^s} \frac{\partial^2 z^k}{\partial x^i \partial x^m} \frac{\partial x^m}{\partial z^q}. \tag{15}$$

现引入记号

$$\Gamma^k_{sq} = -\frac{\partial x^i}{\partial z^s} \frac{\partial x^m}{\partial z^q} \frac{\partial^2 z^k}{\partial x^i \partial x^m}. \tag{16}$$

公式 (15) 便有了最终的形式

$$\widetilde{T}^k_{;q} = \frac{\partial \widetilde{T}^k}{\partial z^q} + \Gamma^k_{sq} \widetilde{T}^s. \tag{17}$$

我们便证明了

定理 28.2. 如果向量场 (T^i) 的梯度在任意坐标变换下如同张量那样变化, 且在欧氏坐标 (x) 下按通常的公式

$$T^i_{;k} = \frac{\partial T^i}{\partial x^k}$$

计算, 则在任意其他的坐标系 (z) 下, 这个梯度的计算公式为

$$\widetilde{T}^k_{;r} = \frac{\partial \widetilde{T}^k}{\partial z^r} + \Gamma^k_{sr}\widetilde{T}^s,$$

其中系数 Γ^k_{sr} 由公式 (16) 定义.

按照类似方法我们可以变换余向量场的表达式:

$$\begin{aligned}
\widetilde{T}_{l;r} &= T_{j;s}\frac{\partial x^j}{\partial z^l}\frac{\partial x^s}{\partial z^r} = \frac{\partial T_j}{\partial x^s}\frac{\partial x^s}{\partial z^r}\frac{\partial x^j}{\partial z^l} = \frac{\partial T_j}{\partial z^r}\frac{\partial x^j}{\partial z^l} \\
&= \frac{\partial}{\partial z^r}\left(\widetilde{T}_k\frac{\partial z^k}{\partial x^j}\right)\frac{\partial x^j}{\partial z^l} = \frac{\partial \widetilde{T}_k}{\partial z^r}\left(\frac{\partial z^k}{\partial x^j}\frac{\partial x^j}{\partial z^l}\right) + \widetilde{T}_k\frac{\partial x^j}{\partial z^l}\frac{\partial}{\partial z^r}\left(\frac{\partial z^k}{\partial x^j}\right) \\
&= \frac{\partial \widetilde{T}_k}{\partial z^r}\delta^k_l + \widetilde{T}_k\frac{\partial x^j}{\partial z^l}\frac{\partial}{\partial z^r}\left(\frac{\partial z^k}{\partial x^j}\right) = \frac{\partial \widetilde{T}_l}{\partial z^r} + \widetilde{T}_k\left(\frac{\partial x^j}{\partial z^l}\frac{\partial x^s}{\partial z^r}\frac{\partial^2 z^k}{\partial x^j \partial x^s}\right) \\
&= \frac{\partial \widetilde{T}_l}{\partial z^r} - \Gamma^k_{lr}\widetilde{T}_k,
\end{aligned}$$

其中 $\Gamma^k_{lr} = -\dfrac{\partial x^j}{\partial z^l}\dfrac{\partial x^s}{\partial z^r}\dfrac{\partial^2 z^k}{\partial x^j \partial x^s}$.

因此, 我们证明了

定理 28.3. 如果余向量场 (T_i) 的梯度在任意坐标变换下如同张量那样变化, 且在欧氏坐标系 (x) 下, 按通常的公式

$$T_{i;k} = \frac{\partial T_i}{\partial x^k}$$

计算, 则在任意坐标系 (z) 下, 这个梯度的计算公式为

$$\widetilde{T}_{i;k} = \frac{\partial \widetilde{T}_i}{\partial z^k} - \Gamma^r_{ik}\widetilde{T}_r, \tag{18}$$

其中数组 Γ^r_{ik} 同于定理 28.2 中对向量场 (T^i) 的表示, 即按公式 (16) 计算.

因为在构造取梯度的运算中, 我们从对变换这样的要求出发即在任意坐标变换 $x = x(z)$ 下其结果如同张量一样; 故而我们得到了对向量和余向量的不同公式.

$$\widetilde{T}_{i;k} = \frac{\partial \widetilde{T}_i}{\partial x^k} - \Gamma^r_{ik}\widetilde{T}_r \quad (\text{对余向量}),$$

$$\widetilde{T}^i_{;k} = \frac{\partial \widetilde{T}^i}{\partial x^k} + \Gamma^i_{rk}\widetilde{T}^r \quad (\text{对向量}).$$

但函数组 Γ^i_{qk} 则是公共的.

我们在这里不对任意 (p, q) 型张量进行计算, 但将不予证明地引入结论:

定理 28.4. 如果 (p,q) 型张量场 $T^{(i)}_{(j)}$ 的梯度 $T^{(i)}_{(j);k}$ 在任意坐标变换下如同张量一样变化, 并且在欧氏坐标系 (x) 下按公式

$$T^{(i)}_{(j);k}=\frac{\partial T^{(i)}_{(j)}}{\partial x^k}$$

计算, 则在任意其他坐标系 $x=x(z)$ 下, 它的计算公式是

$$\widetilde{T}^{(k)}_{(l);r}=\frac{\partial \widetilde{T}^{(k)}_{(l)}}{\partial x^r}+\sum_{s=1}^{p}\widetilde{T}^{k_1\cdots(k_s\to i)\cdots k_p}_{l_1\cdots l_q}\Gamma^{k_s}_{ir}-\sum_{s=1}^{q}\widetilde{T}^{k_1\cdots k_p}_{l_1\cdots(l_s\to i)\cdots l_q}\Gamma^{i}_{l_s r}\qquad(19)$$

(记号 $k_1\cdots(k_s\to i)\cdots k_p$ 表示, 在组 $k_1\cdots k_p$ 中应取 i 替换 k_s), 其中函数组 Γ^{p}_{kq} 按同一公式 (16) 计算.

例如, 对于 2 阶张量我们有

$$\widetilde{T}^{i}_{j;k}=\frac{\partial \widetilde{T}^{i}_{j}}{\partial x^k}+\widetilde{T}^{p}_{j}\Gamma^{i}_{pk}-\widetilde{T}^{i}_{p}\Gamma^{p}_{jk},$$

$$\widetilde{T}_{ij;k}=\frac{\partial \widetilde{T}_{ij}}{\partial x^k}-\widetilde{T}_{pj}\Gamma^{p}_{ik}-\widetilde{T}_{ip}\Gamma^{p}_{jk},$$

$$\widetilde{T}^{ij}_{;k}=\frac{\partial \widetilde{T}^{ij}}{\partial x^k}+\widetilde{T}^{pj}\Gamma^{i}_{pk}+\widetilde{T}^{ip}\Gamma^{j}_{pk}.$$

张量 $T^{(i)}_{(j)}$ 的梯度以 $T^{(i)}_{(j);k}$ 来表示, 其中 $(i)=i_1\cdots i_p,(j)=j_1\cdots j_q$.

应该强调, 我们所引进的运算与欧几里得几何有着本质的联系. 关键之处在于我们用了下面的两个要求来定义这个运算:

a) 运算得到的结果是个张量;

b) 在欧氏坐标下它按照通常的公式计算:

$$T^{(i)}_{(j);k}=\frac{\partial T^{(i)}_{(j)}}{\partial x^k}.$$

(以现在这个运算的观点看, 也可以说, 那样的坐标使在其下的任意张量的梯度可以按上述公式计算时就称此坐标为欧几里得的.)

还要解释函数组 $\Gamma^{k}_{ij}=\Gamma^{k}_{ij}(z)$ 在坐标变换 $z^i=z^i(z'),i=1,\cdots n$ 下是如何变化的.

如果给出了欧氏坐标

$$(x^1,\cdots,x^n),\quad x^i=x^i(z)=x^i(z(z')),$$

则按公式 (16) 我们令

$$\Gamma^{k}_{pq}=-\frac{\partial x^i}{\partial z^p}\frac{\partial x^j}{\partial z^q}\frac{\partial^2 z^k}{\partial x^i\partial x^j}=\frac{\partial^2 x^m}{\partial z^p\partial z^q}\frac{\partial z^k}{\partial x^m}.\qquad(20)$$

在坐标系 (z') 中我们有

$$\Gamma^{k'}_{p'q'} = -\frac{\partial x^i}{\partial z^{p'}}\frac{\partial x^j}{\partial z^{q'}}\frac{\partial^2 z^{k'}}{\partial x^i \partial x^j} = \frac{\partial^2 x^m}{\partial z^{p'}\partial z^{q'}}\cdot\frac{\partial z^{k'}}{\partial x^m}. \tag{21}$$

由公式 (20) 和 (21) 得到了

$$\begin{aligned}\Gamma^k_{pq}\frac{\partial z^p}{\partial z^{p'}}\frac{\partial z^q}{\partial z^{q'}} &= -\frac{\partial^2 z^k}{\partial x^i\partial x^j}\frac{\partial x^i}{\partial z^p}\frac{\partial x^j}{\partial z^q}\frac{\partial z^p}{\partial z^{p'}}\frac{\partial z^q}{\partial z^{q'}}\\ &= -\frac{\partial^2 z^k}{\partial x^i\partial x^j}\frac{\partial x^i}{\partial z^{p'}}\frac{\partial x^j}{\partial z^{q'}} = -\frac{\partial^2 z^k}{\partial z^{p'}\partial z^{q'}} + \frac{\partial z^k}{\partial x^j}\frac{\partial^2 x^j}{\partial z^{p'}\partial z^{q'}},\end{aligned}$$

这里我们用了等式

$$\begin{aligned}\frac{\partial^2[z^k(x(z'))]}{\partial z^{p'}\partial z^{q'}} &= \frac{\partial}{\partial z^{p'}}\left(\frac{\partial z^k}{\partial x^i}\frac{\partial x^i}{\partial z^{q'}}\right)\\ &= \frac{\partial^2 x^i}{\partial z^{p'}\partial z^{q'}}\frac{\partial z^k}{\partial x^i} + \frac{\partial x^i}{\partial z^{p'}}\frac{\partial^2 z^k}{\partial x^i\partial x^j}\frac{\partial x^j}{\partial z^{q'}}.\end{aligned}$$

因此

$$\Gamma^k_{pq}\frac{\partial z^p}{\partial z^{p'}}\frac{\partial z^q}{\partial z^{q'}} + \frac{\partial^2 z^k}{\partial z^{p'}\partial z^{q'}} = \frac{\partial z^k}{\partial x^j}\frac{\partial^2 x^j}{\partial z^{p'}\partial z^{q'}}.$$

由此得到等式

$$\begin{aligned}\frac{\partial z^{k'}}{\partial z^k}\left(\Gamma^k_{pq}\frac{\partial z^p}{\partial z^{p'}}\frac{\partial z^q}{\partial z^{q'}} + \frac{\partial^2 z^k}{\partial z^{p'}\partial z^{q'}}\right) &= \frac{\partial z^{k'}}{\partial z^k}\frac{\partial z^k}{\partial x^j}\frac{\partial^2 x^j}{\partial z^{p'}\partial z^{q'}}\\ &= \frac{\partial^2 x^j}{\partial z^{p'}\partial z^{q'}}\frac{\partial z^{k'}}{\partial x^j} = \Gamma^{k'}_{p'q'}.\end{aligned}$$

最后得到了变换公式

$$\Gamma^{k'}_{p'q'} = \frac{\partial z^{k'}}{\partial z^k}\left(\Gamma^k_{pq}\frac{\partial z^p}{\partial z^{p'}}\frac{\partial z^q}{\partial z^{q'}} + \frac{\partial^2 z^k}{\partial z^{p'}\partial z^{q'}}\right). \tag{22}$$

对于发展更加广泛的张量的共变微分概念而言，我们将利用上面所得的结果作为其理想的基础；这个新概念已不再与原始的欧氏坐标系统有任何关联了.

定义 28.1. 如果在任意坐标系 $z^1,\cdots,z^n$ 中给出了函数组 $\Gamma^k_{pq}(z)$，使它在坐标变换 $z=z(z')$ 下按公式 (22) 变换，则说给出了任意类型的张量的共变微分运算 (取梯度的运算). 称 $\Gamma^k_{pq}(z)$ 为克氏(克里斯托费尔)符号.

对于向量和余向量我们的共变微分运算 (梯度) 的定义公式为

$$T^i_{;k} = \frac{\partial T^i}{\partial x^k} + \Gamma^i_{jk}T^j,$$

$$T_{i;k} = \frac{\partial T_i}{\partial x^k} - \Gamma^j_{ik}T_j,$$

而对一般的张量则是公式 (19). 分量 Γ^k_{ij} 的变换规律 (22) 原本出于要求张量的梯度重新为张量 (虽然 Γ^k_{ij} 并不构成张量). 这个事实将在下一小节中加以证明.

注 1　常常称共变微分 (梯度) 运算为微分 - 几何联络.

注 2　如果存在坐标 $x^1,\cdots,x^n$ 使得 $\Gamma_{ij}^k=0$ (或者在此坐标中 $T_{(j);k}^{(i)}=\dfrac{\partial T_{(j)}^{(i)}}{\partial x^k}$). 也称这样的坐标为欧几里得的.

注 3　共变微分运算常常以符号 ∇ 表示:

$$\nabla_k T_{(j)}^{(i)}=T_{(j);k}^{(i)}.$$

2. 任意阶张量的共变微分

我们已定义了向量 (余向量) 场的共变微分为一个运算, 它在任一个坐标系 $x^1,\cdots,x^n$ 中被写成公式

$$T_{;j}^i=\frac{\partial T^i}{\partial x^j}+\Gamma_{kj}^i T^k\ (\text{向量场}),\tag{23}$$

$$T_{i;j}=\frac{\partial T_i}{\partial x^j}-\Gamma_{ij}^k T_k\ (\text{余向量}).\tag{24}$$

这里的 Γ_{kj}^i 为在所给坐标系中的某个函数, 它在坐标变换下按规律 (22) 变化 (克氏符号).

反过来, 设在每个坐标系 $x^1,\cdots,x^n$ 中给出了函数 $\Gamma_{kj}^i(x)$. 我们以 (23),(24) 给出向量和余向量场的共变微分. 我们将证明, 对于 Γ_{kj}^i 在坐标变换 $x^i=x^i(x^{1'},\cdots,x^{n'})$ 下的变换规律 (22) 被下面的要求条件决定, 即共变微分的运算结果是个张量. 成立下面定理.

定理 28.5.　在坐标变换 $x^i=x^i(x^{1'},\cdots,x^{n'})$ 下函数 Γ_{kj}^i 的变化公式为

$$\Gamma_{k'j'}^{i'}=\Gamma_{kj}^i\frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^k}{\partial x^{k'}}\frac{\partial x^j}{\partial x^{j'}}+\frac{\partial x^{i'}}{\partial x^i}\frac{\partial^2 x^i}{\partial x^{k'}\partial x^{j'}}.\tag{25}$$

证明　因为表达式

$$\frac{\partial T^i}{\partial x^j}+\Gamma_{kj}^i T^k=T_{;j}^i$$

为张量, 则有 (利用等式 $T^i=T^{i'}\dfrac{\partial x^i}{\partial x^{i'}}$)

$$\begin{aligned}
T_{;j'}^{i'}&=\frac{\partial T^{i'}}{\partial x^{j'}}+\Gamma_{k'j'}^{i'}T^{k'}=\left(\frac{\partial T^i}{\partial x^j}+\Gamma_{kj}^i T^k\right)\frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^j}{\partial x^{j'}}\\
&=\frac{\partial x^{i'}}{\partial x^i}\frac{\partial T^i}{\partial x^{j'}}+\Gamma_{kj}^i T^k\frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^j}{\partial x^{j'}}\\
&=\frac{\partial x^{i'}}{\partial x^i}\frac{\partial}{\partial x^{j'}}\left(T^{k'}\frac{\partial x^i}{\partial x^{k'}}\right)+\Gamma_{kj}^i T^{k'}\frac{\partial x^k}{\partial x^{k'}}\frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^j}{\partial x^{j'}}\\
&=\underbrace{\frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^i}{\partial x^{k'}}}_{\delta_{k'}^{i'}}\frac{\partial T^{k'}}{\partial x^{j'}}+T^{k'}\frac{\partial x^{i'}}{\partial x^i}\frac{\partial^2 x^i}{\partial x^{k'}\partial x^{j'}}+\Gamma_{kj}^i T^{k'}\frac{\partial x^k}{\partial x^{k'}}\frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^j}{\partial x^{j'}}
\end{aligned}$$

$$= \frac{\partial T^{i'}}{\partial x^{j'}} + \left(\Gamma^i_{kj} \frac{\partial x^k}{\partial x^{k'}} \frac{\partial x^j}{\partial x^{j'}} \frac{\partial x^{i'}}{\partial x^i} + \frac{\partial x^{i'}}{\partial x^i} \frac{\partial^2 x^i}{\partial x^{k'} \partial x^{j'}} \right) T^{k'}.$$

由此得到所要的论断. □

推论 1 符号 Γ^i_{kj} 只在线性的或者仿射的坐标变换 $x^i = x^i(x^{1'}, \cdots, x^{n'})$, $\frac{\partial^2 x^i}{\partial x^{k'} \partial x^{j'}} \equiv 0$ 下同于张量的变化, 其中恒等式对所有的 i, j', k' 成立.

推论 2 交错表达式

$$T^i_{kj} = \Gamma^i_{kj} - \Gamma^i_{jk} = \Gamma^i_{[kj]} \tag{26}$$

构成一个张量 (挠率张量).

证明 由公式 (25) 可清楚看到, 在指标 k' 和 j' 置换时 $\frac{\partial x^{i'}}{\partial x^i} \frac{\partial^2 x^i}{\partial x^{k'} \partial x^{j'}}$ 这一项并无变化. 因此 $T^{i'}_{k'j'}$ 的变化规律将完全不包含这一项, 从而为张量. □

根据推论 2 的结论, 我们引进定义:

定义 28.2. 如果挠率张量 $T^i_{kj} = \Gamma^i_{[kj]}$ 恒等于零或者 $\Gamma^i_{kj} = \Gamma^i_{jk}$, 则称联络 Γ^i_{kj} 为对称的.

例 28.1. 如果存在欧氏坐标 $x^1, \cdots, x^n$ 即表示 $\Gamma^i_{kj} \equiv 0$, 则挠率张量等于零, 因而欧氏联络是对称的. 在另一坐标 $x^{1'}, \cdots, x^{n'}, x^i = x^i(x')$, 符号 $\Gamma^{i'}_{k'j'}$ 有形状

$$\Gamma^{i'}_{k'j'} = \frac{\partial x^{i'}}{\partial x^i} \frac{\partial^2 x^i}{\partial x^{k'} \partial x^{j'}}.$$

这个表达式对 k', j' 对称.

现在转向具任意阶的张量. 任意阶张量的共变微分被下面的要求唯一确定:

a) 共变微分是个线性算子;

b) 零阶张量 (函数) 的共变微分 (共变导数) 为通常的导数

$$\nabla_k f = \frac{\partial f}{\partial x^k}; \tag{27}$$

c) 向量 (余向量) 场的共变导数由公式 (23) 和 (24) 给出.

d) 张量乘积的共变微分按乘积的微分公式计算:

$$\nabla_k T^{(i)(j)}_{(p)(q)} = (\nabla_k R^{(i)}_{(p)}) S^{(j)}_{(q)} + R^{(i)}_{(p)} (\nabla_k S^{(j)}_{(q)}), \tag{28}$$

其中 $T^{(i)(j)}_{(p)(q)} = R^{(i)}_{(p)} S^{(j)}_{(q)}$ 为张量乘积.

作为基本的例子我们考虑 2 阶张量 T^{ij}, T^i_j, T_{ij}. 有

定理 28.6. 在满足上面所列的共变微分的条件时, 二阶张量的共变微分由下面公式给出:

$$\nabla_k T^{ij} = \frac{\partial T^{ij}}{\partial x^k} + \Gamma^i_{lk} T^{lj} + \Gamma^j_{lk} T^{il}, \tag{29}$$

$$\nabla_k T^i_j = \frac{\partial T^i_j}{\partial x^k} + \Gamma^i_{lk} T^l_j - \Gamma^l_{jk} T^i_l, \tag{30}$$

$$\nabla_k T_{ij} = \frac{\partial T_{ij}}{\partial x^k} - \Gamma^l_{ik} T_{lj} - \Gamma^l_{jk} T_{il}. \tag{31}$$

对于 (p,q) 型张量 $T^{(i)}_{(j)}, (i) = i_1 \cdots i_p, (j) = j_1 \cdots j_q$ 的共变微分按下面公式计算:

$$\nabla_k T^{(i)}_{(j)} = \frac{\partial T^{(i)}_{(j)}}{\partial x^k} - T^{(i)}_{jj_2\cdots j_q} \Gamma^j_{j_1 k} - \cdots - T^{(i)}_{j_1\cdots j} \Gamma^j_{j_q k} +$$

$$T^{ii_2\cdots i_p}_{(j)} \Gamma^{i_1}_{i_k} + \cdots + T^{i_1\cdots i}_{(j)} \Gamma^{i_p}_{ik}. \tag{32}$$

证明 我们对 (0,2) 型张量 T_{ij} 进行证明, 其余的类型完全类似.

设 $e_1, \cdots, e_n$ 为向量场的基;$e^1, \cdots, e^n$ 为余向量场的对偶基:$\langle e^i, e_j \rangle = \delta^i_j$; 这些张量的分量为

$$(e_i)^j = \delta^j_i = (e^j)_i.$$

因此由公式 (23),(24) 我们得到

$$\nabla_k e_i = \Gamma^j_{ik} e_j, \tag{33}$$

$$\nabla_k e^i = -\Gamma^i_{jk} e^j \tag{34}$$

(这些公式可以作为 Γ^i_{jk} 的定义).

具分量 $T^{i_1\cdots i_p}_{j_1\cdots j_q}$ 的任意张量 T 有形式

$$T = T^{i_1\cdots i_p}_{j_1\cdots j_q} e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes e^{j_1} \otimes \cdots \otimes e^{j_q}.$$

特别,(0,2) 型张量的形式为 $T = T_{ij} e^i \otimes e^j$, 利用运算 ∇_k 的性质于是有

$$\begin{aligned}
\nabla_k(T) &= \nabla_k(T_{ij}) e^i \otimes e^j + T_{ij} \nabla_k e^i \otimes e^j + T_{ij} e^i \otimes \nabla_k e^j \\
&= \frac{\partial T_{ij}}{\partial x^k} e^i \otimes e^j - T_{ij} \Gamma^j_{lk} e^l \otimes e^j - T_{ij} e^i \otimes \Gamma^j_{lk} e^l \\
&= \left(\frac{\partial T_{ij}}{\partial x^k} - \Gamma^l_{ik} T_{lj} - \Gamma^l_{jk} T_{il} \right) e^i \otimes e^j.
\end{aligned}$$

因此张量 $\nabla_k T$ 的分量为

$$\frac{\partial T_{ij}}{\partial x^k} - \Gamma^l_{ik} T_{lj} - \Gamma^l_{jk} T_{il},$$

然而按定义这正是 $T_{ij;k}$. 证完. □

注 如果 T^i 为向量场,T_j 为余向量场, 则定义了标量场 T^iT_i (张量乘积的迹). 根据所要求的条件 a)—d) 有公式

$$\begin{aligned}\frac{\partial}{\partial x^k}(T^iT_i) &= (\nabla_k T^i)T_i + T^i(\nabla_k T_i) \\ &= \left(\frac{\partial T^i}{\partial x^k} + \Gamma^i_{jk}T^j\right)T_i + \left(\frac{\partial T_i}{\partial x^k} - \Gamma^j_{ik}T_j\right)T^i \\ &= \frac{\partial}{\partial x^k}(T^iT_i) + \underbrace{\Gamma^i_{jk}T^jT_i - \Gamma^j_{ik}T_jT^i}_{0}.\end{aligned}$$

由此公式我们看到余向量场的共变微分的分量 Γ^j_{ik} 应该与向量场微分的分量的符号相反 (但模相同), 以使对标量 T^iT_i 的公式 $\dfrac{\partial}{\partial x^k}(T^iT_i) = (\nabla_k T^i)T_i + T^i(\nabla_k T_i)$ 成立.

习题

28.1. 设 $\xi = \dot\gamma$ 为曲线 $\gamma = \gamma(t)$ 的速度向量. 令 $\nabla_{\dot\gamma}\eta = \xi^i\nabla_i\eta$, 它定义了向量场 η 沿曲线 γ 的共变导数. 证明场 $\nabla_{\dot\gamma}\eta$ 只依赖于场 η 在曲线 γ 上的值.

§29. 共变微分和度量

1. 向量场的平行移动

如果 $\xi = (\xi^k)$ 为在点 P 处的任一个向量, 则对任意 (p,q) 型张量 $T^{(i)}_{(j)}$ 定义它的沿 ξ 的方向导数为

$$\nabla_\xi T^{(i)}_{(j)} = \xi^k \nabla_k T^{(i)}_{(j)}. \tag{1}$$

它是个张量, 并在此点仍为 (p,q) 型. 对于标量 (零阶张量) 算子 $\nabla_\xi f$ 为

$$\nabla_\xi f = \xi^k \frac{\partial f}{\partial x^k} \equiv \partial_\xi f, \tag{2}$$

即与沿方向 (在已知点上) ξ 的函数 f 的导数. 回忆 §23, 如果我们沿空间曲线

$$x^i = x^i(t), \quad i = 1, \cdots, n$$

运动, 并且如果函数 f 沿此曲线的速度向量方向的导数等于零, 则此函数沿曲线不改变: 如果 $\xi^k\dfrac{\partial f}{\partial x^k} = \dfrac{d}{dt}f(x^1(t),\cdots,x^n(t)) \equiv 0$, 其中 $\xi^k = \dfrac{dx^k}{dt}$ 为速度向量, 于是

$$f(x^1(t),\cdots,x^n(t)) = \text{常数}.$$

是否对向量场甚至张量场成立类似的结果? 这个问题目前并没有意义, 因为向量 (或张量) 在不同的坐标系中有不同的分量; 因此不能比较在空间中不同点给出的两个向量或张量. 这样的比较至少需要空间中一些附加的几何结构; 这个附加的结构就是共变微分 (联络).

设在空间中给出了坐标 $(x^1,\cdots,x^n)$, 共变微分 Γ^i_{kj} 和任一条曲线 $x^i(t), a\leqslant t\leqslant b$.

定义 29.1. 我们说向量 (张量) 场 T 沿曲线段 $x^i=x^i(t), a\leqslant t\leqslant b$ 为共变常向量或平行, 是指这个场 T 在曲线上每点的沿曲线的速度向量的方向共变导数等于零:

$$\nabla_\xi T=\xi^k\nabla_k T=0,\quad a\leqslant t\leqslant b, \tag{3}$$

$$\xi^k=\frac{dx^k}{dt}.$$

对向量场我们有

$$\nabla_\xi T^i=\xi^k\nabla_k T^i=\xi^k\left(\frac{\partial T^i}{\partial x^k}+\Gamma^i_{jk}T^j\right)=0. \tag{4}$$

我们要着重指出, 一般说来平行性的概念依赖于曲线. 例外情形只出现在欧氏几何中: 在欧氏坐标 $(x^1,\cdots,x^n)$ 中我们定义平行向量场是作为在这些欧氏坐标下具有常值分量的场. 这些场显然沿任意曲线都是平行的. 由于共变微分的结果与坐标的选取无关, 尽管在新坐标下它们的分量依赖于点, 这个场将仍然在任意坐标系中平行.

我们将看到, 在不同点的向量的平行性的概念既与共变微分 (微分 几何联络) 有关, 也与联结这两点的路径有关. 为了把我们所引进的几何表示与最基本的中学教材内容联系起来; 我们回忆那个所谓的欧几里得第五公设:"如果在点 P 给出了一条直线, 则通过任意一点 Q 只能引一条平行于它的直线". 对我们来说, 这样来理解这个公设是方便的 (可能并非完全是形式的): 在欧氏几何中, 如果在点 P 给出了向量 $(T^i)_P$, 则在任意点 Q 存在, 并且唯一的一个平行于它的向量 $(T^i_{\|})_Q$.

这里恰好可以提出这样的问题: 这些在不同点 P 和 Q 上附着的广义的平行向量是什么? 按定义, 向量总是附着在已知点上的 (任意张量也一样).

我们来定义向量 T^i 从点 $P=(x_0^1,\cdots,x_0^n)$ 到点 $Q=(x_1^1,\cdots,x_1^n)$ 沿曲线 $x^i(t)$ 平行移动的重要概念, 其中 $x^i(0)=x_0^i, x^i(1)=x_1^i$.

定义 29.2. 设 T_P^i 为在点 $P=(x_0^1,\cdots,x_0^n)$ 的向量,$x^i=x^i(t), 0\leqslant t\leqslant 1$ 为一曲线, 由点 P 到点 $Q=(x_1^1,\cdots,x_1^n)$. 如果在曲线上所有点上给出了向量场 T^i, 它沿此曲线平行: $\dfrac{dx^i}{dt}\nabla_k T^i=0$ 对所有 $0\leqslant t\leqslant 1$ 成立, 则称向量 T_P^i 沿此曲线从点 P 平移到点 Q. 当 $t=0$ 时, 向量场 T^i 在点 P 应该与原来的向量 T_P^i 重合. 当 $t=1$ 时向量场在点 Q 为向量 T_Q^i, 称它为 T_P^i 沿所给曲线 $x^i=x^i(t)$ 从 P 到 Q 平移的结果.

在坐标 $x^1, \cdots, x^n$ 中我们得到

$$\frac{dx^k}{dt}\nabla_k T^i = \frac{\partial T^i}{\partial x^k}\frac{dx^k}{dt} + T^j \Gamma^i_{jk}\frac{dx^k}{dt} - \frac{dT^i}{dt} + \left(\frac{dx^k}{dt}\Gamma^i_{jk}\right)T^j = 0. \tag{5}$$

这是平移的方程. 初始条件 (当 $t=0$) 为

$$T^i(0) = T^i_P, \quad i = 1, \cdots, n. \tag{6}$$

方程 (5) 对 T^i 是线性的. 由微分方程解的存在性和唯一性定理以及解的拓展性定理, 对于任意光滑曲线我们有如下结果:

定理 29.1. 沿任意一条固定的光滑曲线存在平行移动, 其结果由初始向量 T^i_P 唯一决定并线性依赖于它.

因此向量沿曲线的平移依赖于联络 Γ^i_{jk}. 如果联络是欧氏的, 即 $\Gamma^i_{jk} \equiv 0$ 在欧氏坐标下成立, 于是我们这时得到了平移的方程 $\dfrac{dT^i}{dt} = 0$.

推论 在欧氏几何中并在欧氏坐标下, 曲线上两个不同点出发的向量如果有相同的分量则沿此曲线是平行的. 在任意坐标中, 如果几何是欧几里得的, 则沿曲线平移向量的结果不依赖于曲线自身.

现在在弯曲的和欧氏几何之间的直观区别已经清楚了: 沿不同曲线同一个向量从点 P 到 Q 的平移, 在曲率存在的情形 (即具非零曲率) 下, 其平移结果得到不同的向量.

关于曲率的数值量度我们转到下一节去讨论.

2. 测地线

我们转而考虑在任意联络下类似于直线的曲线. 称这些曲线为测地线.

定义 29.3. 如果曲线 $x^i = x^i(t)$ 的速度向量 $T^i = \dfrac{dx^i}{dt}$ 沿自身平行:

$$\nabla_T(T) = 0, \tag{7}$$

则称此曲线是测地线.

在坐标表示下有

$$\begin{aligned} 0 &= \nabla_T(T)^j = \frac{dx^i}{dt}\nabla_i\left(\frac{dx^j}{dt}\right) = \frac{dx^i}{dt}\left[\frac{\partial}{\partial x^i}\left(\frac{dx^j}{dt}\right) + \Gamma^j_{ki}\frac{dx^k}{dt}\right] \\ &= \frac{d^2x^j}{dt^2} + \Gamma^j_{ki}\frac{dx^k}{dt}\frac{dx^i}{dt}. \end{aligned}$$

因此, 我们得到了测地方程

$$\frac{d^2x^j}{dt^2} + \Gamma^j_{ki}\frac{dx^k}{dt}\frac{dx^i}{dt} = 0, \quad j = 1, \cdots, n. \tag{8}$$

如果 $\Gamma^j_{ki}=0$, 则通常的直线便是这个方程的解, 这应该就是欧氏几何中的情形.

对于任意的联络, 方程 (8) 为二阶微分方程组. 在点 $(x^1_0,\cdots,x^n_0)$ 的邻域中, 这个方程在初值条件

$$x^j|_{t=0}=x^j_0,\quad \left.\frac{dx^j}{dt}\right|_{t=0}=\dot{x}^j_0;\quad j=1,\cdots,n \tag{9}$$

下存在唯一的解, 其中 $x^j_0,\dot{x}^j_0$ 任意 (由解的存在和唯一性定理得到). 故而成立下面的定理.

定理 29.2. 在任意点 P 的某个邻域中, 对此点上的任意向量 T^i_P, 存在联络 (Γ^i_{jk}) 的唯一的测地线, 它由 P 点出发并且初始速度向量为 T^i_P.

注 由方程 (8) 看出, 由联络给出的测地线只依赖于联络的 "对称部分" $\Gamma^i_{(jk)}=\Gamma^i_{jk}+\Gamma^i_{kj}$.

3. 与度量相容的联络

在本书中我们有过关于欧氏坐标的两个定义:

1) 如果在坐标 $x^1,\cdots,x^n$ 下度量 g_{ij} 有欧几里得形式

$$g_{ij}=\delta_{ij}, \tag{10}$$

则说此坐标是欧几里得的.

2) 如果在这个坐标下联络的分量 Γ^k_{ij} 为零:

$$\Gamma^k_{ij}\equiv 0,\quad i,j,k=1,\cdots,n, \tag{11}$$

则说此坐标是欧几里得的.(更一般地, 对于非对称联络有 $\Gamma^k_{ij}=-\Gamma^k_{ji}$).

这两个欧氏坐标的定义之间有什么相互关联?

应可立即看到, 联络概念和黎曼度量概念之间没有联系. 在所考虑的空间区域中这是两个独立的结构, 故定义 1) 和 2) 是两个不同概念的定义.

但是存在使度量与联络匹配的方法, 这时定义 1) 和 2) 给出同样的结果, 它们之间仅差一个仿射变换.

定义 29.4. 如果度量张量的共变导数恒等于零:

$$\nabla_k g_{ij}\equiv 0,\quad k,i,j=1,\cdots,n, \tag{12}$$

则称联络 Γ^k_{ij} 与度量 g_{ij} 相容.

我们指出与度量相容的联络的两个性质.

1) 如果联络与度量相容, 则降低任一张量指标的运算与共变微分可交换.

证明 由运算 ∇_k 的性质 (莱布尼茨公式) 和等式 (12) 得到公式

$$\nabla_k(g_{lm}T^{(i)}_{(j)})=g_{lm}(\nabla_k T^{(i)}_{(j)}),$$

其中张量 $T_{(j)}^{(i)}$ 为任意的 (p,q) 型. 因为共变微分算子是线性的, 故由此得到所要的论断.

2) 如果向量场 $T^i(t)$ 和 $S^i(t)$ 沿曲线 $x^i = x^i(t)$ 都平行, 则它们的内积沿此曲线为常数.

证明 我们将证明 $\dfrac{d}{dt}(T,S) = \dfrac{d}{dt}(g_{ij}T^iS^j) = 0$. 我们有

$$\begin{aligned}\frac{d}{dt}(g_{ij}T^iS^j) &= \frac{dx^k}{dt}\nabla_k(g_{ij}T^iS^j) = \frac{dx^k}{dt}g_{ij}\nabla_k(T^iS^j)\\ &= g_{ij}\left(\frac{dx^k}{dt}\nabla_kT^i\right)S^j + g_{ij}T^i\left(\frac{dx^k}{dt}\nabla_kS^j\right) = 0\end{aligned}$$ □

换句话说, 如果联络与度量相容, 则沿已知曲线从点 P 到点 Q 的向量的平行移动是在点 P 的切空间到点 Q 的切空间的正交变换

如何描述与已知度量相容的联络? 有重要定理:

定理 29.3. 如果度量 g_{ij} 非退化 (即 $g = \det(g_{ij}) \neq 0$), 则存在唯一的联络, 它为对称且与此度量相容. 在任一坐标系 $(x^1,\cdots,x^n)$ 下, 这个联络由公式

$$\Gamma_{ij}^k = \frac{1}{2}g^{kl}\left(\frac{\partial g_{lj}}{\partial x^i} + \frac{\partial g_{il}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^l}\right) \tag{13}$$

给出 (克里斯托费尔公式).

证明 由定义我们有

$$\begin{aligned}&\Gamma_{ij}^k = \Gamma_{ji}^k,\\ &\nabla_k g_{ij} = \frac{\partial g_{ij}}{\partial x^k} - \Gamma_{ik}^l g_{lj} - \Gamma_{jk}^l g_{il} = 0.\end{aligned} \tag{14}$$

我们想对后面一个方程解出 Γ_{ij}^k. 由降标的定义, 我们有

$$\Gamma_{k,ij} = g_{kl}\Gamma_{ij}^l.$$

方程 (14) 有形式

$$\Gamma_{i,jk} + \Gamma_{j,ik} = \frac{\partial g_{ij}}{\partial x^k}.$$

其中 $\Gamma_{i,jk} = \Gamma_{i,kj}, \Gamma_{j,ik} = \Gamma_{j,ki}$. 循环置换指标 i,j,k, 我们得到

$$\Gamma_{i,jk} + \Gamma_{j,ik} = \frac{\partial g_{ij}}{\partial x^k},$$

$$\Gamma_{j,ki} + \Gamma_{k,ji} = \frac{\partial g_{kj}}{\partial x^i},$$

$$\Gamma_{i,kj} + \Gamma_{k,ij} = \frac{\partial g_{ik}}{\partial x^j}.$$

如果 $(a), (b), (c)$ 为这些公式的左端, 则由于联络的对称性成立等式 $(b)+(c)-(a)=2\Gamma_{k,ij}$. 故而

$$\Gamma_{k,ij}=\frac{1}{2}\left(\frac{\partial g_{kj}}{\partial x^i}+\frac{\partial g_{ik}}{\partial x^j}-\frac{\partial g_{ji}}{\partial x^k}\right)=g_{kl}\Gamma^l_{ij}.$$

提升指标 k, 我们得到所要的公式. 定理得证. □

推论 如果选择坐标使得在所给点上 g_{ij} 的每个一阶导数为零, 则在这个点上克氏符号 Γ^k_{ij} (对于对称的, 与度量相容的联络) 等于零.

例 29.1. 考虑位于三维欧氏空间中的曲面; 空间的坐标为 $x^1=x, x^2=y, x^3=z$(欧几里得的), 曲面为

$$x^1=x^1(z^1,z^2),\quad x^2=x^2(z^1,z^2),\quad x^3=x^3(z^1,z^2).$$

像我们在 §8 中已经做过的那样, 我们假定 x^3 轴垂直于曲面在点 P 的切平面, 而 x^1, x^2 轴为它的参数. 我们在点 P 附近选取作为参数的 x^1, x^2; 于是此曲面由方程

$$x^3=f(z^1,z^2),\quad z^1=u=x^1,\quad z^2=v=x^2,$$

从而 $z^1=x^1, z^2=x^2$; 另外, 因为 x^3 轴在点 $P=(0,0)$ 垂直于曲面, 我们在点 $P=(0,0)$ 得到

$$\left.\frac{\partial f}{\partial z^1}\right|_{(0,0)}=\left.\frac{\partial f}{\partial z^2}\right|_{(0,0)}=0\quad\text{或}\quad \left.\operatorname{grad} f\right|_{(0,0)}=0.$$

对于度量 g_{ij} 我们有 (§7.3)

$$g_{ij}=\delta_{ij}+\frac{\partial f}{\partial z^i}\frac{\partial f}{\partial z^j}.$$

在点 P 因 $\dfrac{\partial f}{\partial z^i}=0$, 故 $g_{ij}=\delta_{ij}$, 并且

$$\frac{\partial g_{ij}}{\partial z^k}=\frac{\partial}{\partial z^k}\left(\frac{\partial f}{\partial z^i}\frac{\partial f}{\partial z^j}\right)=\frac{\partial^2 f}{\partial z^i\partial z^k}\frac{\partial f}{\partial z^j}+\frac{\partial^2 f}{\partial z^j\partial z^k}\frac{\partial f}{\partial z^i}=0.$$

因此在点 P 的这些坐标中所有的符号 Γ^q_{ik} 等于 0 $(q,i,k=1,2)$.

例 29.2. 对向量场 $T=(T^i)$ 定义其发散量为

$$\operatorname{div} T=\nabla_i T^i=T^i_{;i}. \tag{15}$$

对于对称联络, 如与黎曼 (伪黎曼) 度量相容, 我们则有 $\nabla_i T^i=\dfrac{\partial T^i}{\partial x^i}+\Gamma^i_{ki}T^k$, 且

$$\begin{aligned}\Gamma^i_{ki}&=\frac{1}{2}g^{il}\left(\frac{\partial g_{lk}}{\partial x^i}+\frac{\partial g_{li}}{\partial x^k}-\frac{\partial g_{ki}}{\partial x^l}\right)=\frac{1}{2}g^{il}\frac{\partial g_{il}}{\partial x^k}=\frac{1}{2g}\frac{\partial g}{\partial x^k}\\&=\frac{\partial}{\partial x^k}\ln(\sqrt{|g|}),\end{aligned}$$

其中 $g=\det(g_{ij})$. 因此,

$$\nabla_i T^i = \frac{\partial T^i}{\partial x^i} + \frac{1}{2g}\frac{\partial g}{\partial x^k}T^k = \frac{1}{\sqrt{|g|}}\frac{\partial}{\partial x^i}(\sqrt{|g|}T^i). \tag{16}$$

结论 向量场 (T^i) 的发散量 $\nabla_i T^i$ 具有通常的形式 $\nabla_i T^i = \dfrac{\partial T^i}{\partial x^i}$ 当且仅当在体积元 $\sqrt{|g|}dx^1\wedge\cdots\wedge dx^n$ 与欧几里得的体积元相合:$\sqrt{|g|}=1$, 其中 $g=\det(g_{ij})$.

现在我们已有了在联络和黎曼度量 g_{ij} 之间的联系 (以共变微分的方式): 任意一个黎曼几何都产生了确定的对称的对张量的微分方式, 按此方式度量自身可以看作为常值.

4. 与复结构相容的联络

考虑在复空间中的区域 D, 空间的坐标为 $z^1,\cdots,z^n, z^k=x^k+iy^k$, 其中 $x^1,\cdots,x^n,y^1,\cdots,y^n$ 为实区域 $D^{\mathbb{R}}$ 中的实坐标. 设在区域 D 中给出了埃尔米特度量

$$ds^2 = h_{i\bar{k}}dz^i d\bar{z}^k, \quad h_{k\bar{i}} = h_{i\bar{k}} \tag{17}$$

(记作 (见 §27), 对应于 $d\bar{z}^k$ 的指标 k, 我们要加上 "杠"). 它在区域 $D^{\mathbb{R}}$ 中定义了一个黎曼度量, 其公式 (见 §11.2) 为

$$ds^2_{\mathbb{R}} = \mathrm{Re}(h_{i\bar{k}})[dx^i dx^k + dy^i dy^k]. \tag{18}$$

由于在上一小节中已说过, 在 $D^{\mathbb{R}}$ 中存在唯一的对称联络, 它与度量 (18) 相容. 一般说来, 这个联络并不与 D 中复结构相容, 这里所说的相容的意义是指向量沿路径的平移不是个酉变换. 有下面的定理.

定理 29.4. 与度量 ds^2 相容的对称联络也与区域 D 的复结构相容当且仅当度量 ds^2 是凯勒的.

证明 回想 (参看 §27) 我们称埃尔米特度量 ds^2 为凯勒是说由公式

$$\Omega = \frac{i}{2}h_{j\bar{k}}dz^j\wedge d\bar{z}^k \tag{19}$$

定义的微分形式为闭:

$$d\Omega = 0 \Leftrightarrow \frac{\partial h_{j\bar{k}}}{\partial z^i} - \frac{\partial h_{i\bar{k}}}{\partial z^j} = 0, \quad \frac{\partial h_{j\bar{k}}}{\partial \bar{z}^i} - \frac{\partial h_{j\bar{i}}}{\partial \bar{z}^k} = 0. \tag{20}$$

在切空间中引进复基

$$\frac{\partial}{\partial z^k} = \frac{1}{2}\left(\frac{\partial}{\partial x^k} - i\frac{\partial}{\partial y^k}\right), \quad \frac{\partial}{\partial \bar{z}^k} = \frac{1}{2}\left(\frac{\partial}{\partial x^k} + i\frac{\partial}{\partial y^k}\right), \quad k=1,\cdots,n.$$

任意向量 ξ 有形式 $\xi=(\xi^k,\xi^{\bar{k}})$:

$$\xi = \xi^k\frac{\partial}{\partial z^k} + \xi^{\bar{k}}\frac{\partial}{\partial \bar{z}^k},$$

如果 ξ 为实向量则 $\xi^{\bar{k}}=\overline{\xi^k}$. 在这组基下, 内积 $ds^2_{\mathbb{R}}$ 由矩阵 $g_{\alpha\beta}$ 给出, $\alpha,\beta=1,\cdots,n,\bar{1},\cdots,\bar{n}$:

$$g_{ij}=g_{\bar{i}\bar{j}}=0;\quad g_{i\bar{j}}=h_{i\bar{j}},\quad g_{\bar{i}j}=\overline{h_{i\bar{j}}},\tag{21}$$

即矩阵 $G=(g_{\alpha\beta})$ 有形式

$$G=\begin{pmatrix}0 & H\\ \bar{H} & 0\end{pmatrix},\quad H=(h_{i\bar{j}}).\tag{22}$$

于是逆矩阵 G^{-1} 有形式

$$G^{-1}=\begin{pmatrix}0 & \bar{H}^{-1}\\ H^{-1} & 0\end{pmatrix}.\tag{23}$$

我们来计算与度量 $g_{\alpha\beta}$ 相容的联络的分量 $\Gamma^{\alpha}_{\beta\gamma}$, 我们要用克里斯托费尔公式. 由度量的形状 $g_{\alpha\beta}$ 我们有

$$\Gamma^i_{jk}=\frac{1}{2}g^{i\bar{m}}\left(\frac{\partial g_{j\bar{m}}}{\partial z^k}+\frac{\partial g_{\bar{m}k}}{\partial z^j}\right);$$
$$\Gamma^{\bar{i}}_{\bar{j}\bar{k}}=\frac{1}{2}g^{m\bar{i}}\left(\frac{\partial g_{\bar{k}m}}{\partial \bar{z}^j}+\frac{\partial g_{m\bar{j}}}{\partial \bar{z}^k}\right)=\overline{\Gamma^i_{jk}}.$$

另外,

$$\Gamma^i_{\bar{j}\bar{k}}=\frac{1}{2}g^{i\bar{m}}\left(\frac{\partial g_{\bar{m}\bar{k}}}{\partial \bar{z}^j}+\frac{\partial g_{\bar{j}\bar{m}}}{\partial \bar{z}^k}-\frac{\partial g_{\bar{j}\bar{k}}}{\partial \bar{z}^m}\right)=0;$$

按相似的理由, 有 $\Gamma^{\bar{i}}_{jk}=0$. 最后,

$$\Gamma^i_{j\bar{k}}=\frac{1}{2}g^{i\bar{m}}\left(\frac{\partial g_{j\bar{m}}}{\partial \bar{z}^k}-\frac{\partial g_{j\bar{k}}}{\partial \bar{z}^m}\right)=\frac{1}{2}g^{i\bar{m}}\left(\frac{\partial h_{j\bar{m}}}{\partial \bar{z}^k}-\frac{\partial h_{j\bar{k}}}{\partial \bar{z}^m}\right)=0$$

是由于凯勒条件, 此表达式为 0. 以相似的方式验证, 凯勒条件保证了分量 $\Gamma^i_{\bar{j}k},\Gamma^{\bar{i}}_{j\bar{k}}$, $\Gamma^{\bar{i}}_{\bar{j}k}$ 都化为零.

因此, 我们已证明在 $\Gamma^{\alpha}_{\beta\gamma}$ 中只有 Γ^i_{jk} 和 $\Gamma^{\bar{i}}_{\bar{j}\bar{k}}$ 不为零, 而且

$$\Gamma^{\bar{i}}_{\bar{j}\bar{k}}=\overline{\Gamma^i_{jk}}.$$

我们考虑向量 $\xi=(\xi^i,\xi^{\bar{i}})$ 在按 $(\delta z^j,\delta\bar{z}^j)$ 的无穷小平移下的变化, 其中 $\delta\bar{z}^j=\overline{\delta z^j}$. 由上面已证过的事实, 我们有

$$\begin{aligned}\xi^i&\to\xi^i-\xi^k\Gamma^i_{kj}\delta z^j,\\ \xi^{\bar{i}}&\to\xi^{\bar{i}}-\xi^{\bar{k}}\Gamma^{\bar{i}}_{\bar{k}\bar{j}}\delta\bar{z}^j.\end{aligned}\tag{24}$$

因此, 如果 $A=(a^i_k)$ 为矩阵

$$a^i_k=\Gamma^i_{kj}\delta z^j,$$

则对实位移 $(\delta z^j, \delta\bar{z}^j)$ (其中 $\delta z^j = \overline{\delta\bar{z}^j}$) 的相应的平移矩阵为

$$\begin{pmatrix} 1-A & 0 \\ 0 & 1-\bar{A}, \end{pmatrix}, \tag{25}$$

它表示了平移的复线性性 (参看 §12.2) □

习题

29.1. 证明联络与度量相容当且仅当对任何向量场 η, ξ_1, ξ_2 成立

$$\partial_\eta\langle\xi_1, \xi_2\rangle = \langle\nabla_\eta\xi_1, \xi_2\rangle + \langle\xi_1, \nabla_\eta\xi_2\rangle.$$

29.2. 证明在向量的无穷小平移 δx^k 下, 它的分量的变化为下面方式 (准确到更高阶的无穷小):

$$\xi^i \to \xi^i - \xi^j \Gamma^i_{jk}\delta x^k + o(|\delta x|).$$

29.3. 用共变导数表示 §23.3 中的形变张量.

29.4. 设在区域 U 上给出了联络. P 为此区域中一个固定点, $T = T_P$ 为 U 在此点的切空间. 我们定义映射 $E: T \to U$: 设 ξ 为 T 中向量. 由点 P 出发以 ξ 为初始速度向量作测地线 $\gamma(t)$. 令 $E(\xi) = \gamma_\xi(1)$. a) 证明映射 E 在 T 的原点的某个邻域中有定义, 并在其中为局部微分同胚. b) 证明在由映射 E 定义的坐标下所有符号 Γ^k_{ij} 在点 P 化为零.

29.5. 点电荷在磁极场中的运动方程为

$$\ddot{r} = a\frac{[r, \dot{r}]}{|r|^3}, \quad a = \text{常数}.$$

证明电荷的轨线为圆锥上的测地线.

29.6. 求在罗巴切夫斯基平面上所有的测地线.

29.7. 证明球面上的测地线为大圆而且也只为大圆.

29.8. 利用测地线证明保持一点不变且保持在这点的标架不变的运动为恒同映射.

29.9. 证明函数

$$z(u, v) = \int\frac{du}{\sqrt{f(u) - a}} \pm \int\frac{dv}{\sqrt{g(v) + a}}$$

的水平曲线为在度量

$$dl^2 = (f(u) + g(v))(du^2 + dv^2), \quad f > 0, g > 0$$

下的测地线.

29.10. 证明内自同构 $X \mapsto AXA^{-1}$ 是保持群单位元不动的 $SO(3,\mathbb{R})$ 上基灵度量下的运动, 其中 $X, A \in SO(3,\mathbb{R})$.

29.11. 对于与度量 g_{ij} 相容的对称联络, 证明下面恒等式的正确性:

a) $g^{kl}\Gamma^i_{kl} = -\dfrac{1}{\sqrt{|g|}}\dfrac{\partial}{\partial x^k}(\sqrt{|g|}g^{ik})$;

b) $\Gamma^i_{ki} = \dfrac{1}{2g}\dfrac{\partial g}{\partial x^k}$.

29.12. 证明在黎曼空间中两个邻近点可以用测地线相连接, 这样的测地线是局部唯一的.

29.13. 设度量为 $dl^2 = g_{rr}dr^2 + r^2d\varphi^2$. 证明由中心出发的线 $\varphi = \varphi_0$ 为测地线.

29.14. 在度量 (g_{ij}) 的 n 维空间中有下面的公式

$$\oint\limits_{\partial V} X^i dS_i = \int\limits_V \nabla_i X^i \sqrt{|g|}dx^1 \wedge \cdots \wedge dx^n,$$

其中

$$dS_i = \frac{1}{(n-1)!}\sqrt{|g|}\varepsilon_{i_1\cdots i_{n-1}i}dx^{i_1} \wedge \cdots \wedge dx^{i_{n-1}},$$

(参考习题 26. 2).

29.15. 设 M 为欧氏空间 $\mathbb{R}^n$ 的曲面,π 为线性算子, 将 $\mathbb{R}^n$ 正交投影到曲面 M 的切空间中;X, Y 为 $\mathbb{R}^n$ 中的向量场, 它们切于平面 M. 证明与诱导在曲面上的度量相容的联络有形式

$$\nabla_X Y = \pi\left(X^k\frac{\partial Y}{\partial x^k}\right).$$

§30. 曲率张量

1. 一般曲率张量

在前一节中我们解释过, 在非欧氏空间中向量的平移依赖于路径. 同时向量 T 沿路径 $x(t)$ 的平移结果由移动的方程

$$\frac{dT^i}{dt} + \Gamma^i_{jk}T^k\frac{dx^j}{dt} = 0 \tag{1}$$

定义, 我们需要解此方程.

极其方便的办法是不去解此方程, 而是决定出联络 (Γ^i_{jk}) 与欧氏联络相差异的局部特征. 这些特征是什么? 如何知道是否存在坐标 $x^1, \cdots, x^n$, 使 $\Gamma^i_{jk} = 0$? (当然, 如果联络为非对称, 则 $T^i_{jk} = \Gamma^i_{jk} - \Gamma^i_{kj}$ 为非零张量; 因此不可能引入坐标使得 $\Gamma^i_{jk} \equiv 0$. 在此情形下我们可以解释坐标为欧几里得表示 $\Gamma^i_{jk} = \dfrac{1}{2}T^i_{jk}$ (即联络 Γ^i_{jk}) 对下指标为反称 (其对称部分 $\Gamma^i_{jk} + \Gamma^i_{kj} \equiv 0$).)

如何解释关于欧氏坐标的存在性问题? 我们对对称联络的这个问题感兴趣. 我们知道在普通分析中的偏导数的重要性质:

$$\frac{\partial}{\partial x^j}\frac{\partial f}{\partial x^i} = \frac{\partial}{\partial x^i}\frac{\partial f}{\partial x^j} = \frac{\partial^2 f}{\partial x^i \partial x^j}.$$

如果此联络允许坐标 $x^1, \cdots, x^n$ 是欧氏的, 则在这坐标下以通常的公式去微分张量:

$$\nabla_k T^{(i)}_{(j)} = \frac{\partial T^{(i)}_{(j)}}{\partial x^k}.$$

因此

$$(\nabla_k\nabla_l - \nabla_l\nabla_k)T^{(i)}_{(j)} = 0$$

或者

$$T^{(i)}_{(j);k;l} = T^{(i)}_{(j);l;k}.$$

因为 $T^{(i)}_{(j);k;l}$ 是张量, 故上面这个性质在任意坐标下都成立. 我们来看一下一般的联络.

对在任意坐标 $x^1, \cdots, x^n$ 下的向量场, 我们有

$$\begin{aligned}\nabla_k\nabla_l T^i &= \nabla_k\left(\frac{\partial T^i}{\partial x^l} + \Gamma^i_{ql}T^q\right)\\ &= \frac{\partial}{\partial x^k}\left(\frac{\partial T^i}{\partial x^l} + \Gamma^i_{ql}T^q\right) + \Gamma^i_{pk}\left(\frac{\partial T^p}{\partial x^l} + \Gamma^p_{ql}T^q\right) - \Gamma^p_{lk}\left(\frac{\partial T^i}{\partial x^p} + \Gamma^i_{qp}T^q\right)\\ &= \frac{\partial^2 T^i}{\partial x^k \partial x^l} + \frac{\partial T^q}{\partial x^k}\Gamma^i_{ql} + \Gamma^i_{pk}\frac{\partial T^p}{\partial x^l} - \Gamma^p_{lk}\frac{\partial T^i}{\partial x^p} + T^q\frac{\partial \Gamma^i_{ql}}{\partial x^k} + \Gamma^i_{pk}\Gamma^p_{ql}T^q - \Gamma^p_{lk}\Gamma^i_{qp}T^q.\end{aligned}$$

构造 $(\nabla_k\nabla_l - \nabla_l\nabla_k)T^i$. 在简化后我们得到在坐标下的表示式

$$\begin{aligned}&(\nabla_k\nabla_l - \nabla_l\nabla_k)T^i\\ &= \left(\frac{\partial \Gamma^i_{ql}}{\partial x^k} - \frac{\partial \Gamma^i_{qk}}{\partial x^l}\right)T^q + \left(\Gamma^i_{pk}\Gamma^p_{ql} - \Gamma^i_{pl}\Gamma^p_{qk}\right)T^q - (\Gamma^p_{lk} - \Gamma^p_{kl})\frac{\partial T^i}{\partial x^p}.\end{aligned}$$

引入记号

$$-R^i_{qkl} = \frac{\partial \Gamma^i_{ql}}{\partial x^k} - \frac{\partial \Gamma^i_{qk}}{\partial x^l} + \Gamma^i_{pk}\Gamma^p_{ql} - \Gamma^i_{pl}\Gamma^p_{qk}. \tag{2}$$

于是得到公式

$$(\nabla_k\nabla_l - \nabla_l\nabla_k)T^i = -R^i_{qkl}T^q + T^p_{kl}\frac{\partial T^i}{\partial x^p}, \tag{3}$$

其中 T^p_{kl} 为挠率张量(参看 §28). 我们发现 R^i_{qkl} 原来是个张量; 称此张量为黎曼张量或黎曼曲率.对于对称联络, 我们有 $T^p_{kl} \equiv 0$. 因此我们有下面定理:

定理 30.1. 对于对称联络和任意的向量场 T, 表达式 $(\nabla_k\nabla_l-\nabla_l\nabla_k)T^i$ 有形式 $-R^i_{qkl}T^q$, 其中 R^i_{qkl} 为黎曼张量, 它由下面公式定义:

$$-R^i_{qkl}=\frac{\partial\Gamma^i_{ql}}{\partial x^k}-\frac{\partial\Gamma^i_{qk}}{\partial x^l}+\Gamma^i_{pk}\Gamma^p_{ql}-\Gamma^i_{pl}\Gamma^p_{qk}.$$

如果此联络为欧几里得的, 则 $R^i_{qkl}\equiv 0$. 在使 $\Gamma^i_{pq}=0$ 的那些点上成立等式

$$-R^i_{qkl}=\frac{\partial\Gamma^i_{ql}}{\partial x^k}-\frac{\partial\Gamma^i_{qk}}{\partial x^l}.$$

由此定理和 R^i_{qkl} 的张量特性得到

推论 如果黎曼张量不为零, 则不可能引进欧氏坐标, 使在此坐标下 $g_{ij}=\delta_{ij}$ 或 $\Gamma^k_{ij}=0$.

注 这个结果可以以其他方式得到. 考虑分量 Γ^k_{ij} 在变换 $x=x(x')$ 下的变化规律, 于是有

$$\Gamma^{k'}_{i'j'}=\frac{\partial x^{k'}}{\partial x^k}\left(\Gamma^k_{ij}\frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^j}{\partial x^{j'}}+\frac{\partial^2x^k}{\partial x^{i'}\partial x^{j'}}\right).$$

设联络为对称: $\Gamma^k_{ij}=\Gamma^k_{ji}$. 我们要去找坐标 x', 使得 $\Gamma^{k'}_{i'j'}\equiv 0$. 对于这样的 $x^i=x^i(x^{1'},\cdots,x^{n'})$, 我们得到方程

$$\frac{\partial^2x^k}{\partial x^{i'}\partial x^{j'}}=-\Gamma^k_{ij}\frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^j}{\partial x^{j'}},\quad \Gamma^k_{ij}=\Gamma^k_{ij}(x). \tag{*}$$

可以解出这些方程吗? 如果它们可解, 则

$$\frac{\partial}{\partial x^{k'}}\left(\frac{\partial^2x^k}{\partial x^{i'}\partial x^{j'}}\right)-\frac{\partial}{\partial x^{i'}}\left(\frac{\partial^2x^k}{\partial x^{k'}\partial x^{j'}}\right)=0.$$

这可化作方程 (*) 右端的一个条件, 它等价于等式

$$R^i_{qkl}\equiv 0.$$

我们以不变的记号引进曲率和挠率的公式. 如果 ξ,η,ζ 为向量场, 则令

$$[T(\xi,\eta)]^i=T^i_{kl}\xi^k\eta^l, \tag{4}$$

$$[R(\xi,\eta)\zeta]^i=R^i_{jkl}\xi^k\eta^l\zeta^j. \tag{5}$$

我们来给出向量场 $T(\xi,\eta),R(\xi,\eta)\zeta$ 通过场 ξ,η,ζ 的显式表达式.

引理 30.1. 对任意向量场 ξ,η,ζ 成立等式

$$T(\xi,\eta)=\nabla_\xi\eta-\nabla_\eta\xi-[\xi,\eta], \tag{6}$$

$$R(\xi,\eta)\zeta=\nabla_\eta\nabla_\xi\zeta-\nabla_\xi\nabla_\eta\zeta+\nabla_{[\xi,\eta]}\zeta, \tag{7}$$

其中 $[\xi,\eta]$ 为向量场的换位子.

证明 首先验证在公式 (6) 和 (7) 中,$T(\xi,\eta)$ 和 $R(\xi,\eta)\zeta$ 线性依赖于向量场 ξ,η,ζ 的坐标. 对于 $T(\xi,\eta)$ 我们有: 如果 $\xi\to f\xi$, 其中 f 为光滑函数, 则

$$\begin{aligned}T(f\xi,\eta) &= \nabla_{f\xi}\eta-\nabla_\eta(f\xi)-[f\xi,\eta]\\&= f[\nabla_\xi\eta-\nabla_\eta\xi-[\xi,\eta]]-(\partial_\eta f)\xi+(\partial_\eta f)\xi\\&= fT(\xi,\eta),\end{aligned}$$

其中 $\partial_\eta f$ 为沿向量场 η 的函数 f 的方向导数. 对于 $R(\xi,\eta)$

$$\begin{aligned}R(f\xi,\eta) &= \nabla_\eta\nabla_{f\xi}-\nabla_{f\xi}\nabla_\eta+\nabla_{[f\xi,\eta]}\\&= f\nabla_\eta\nabla_\xi+(\partial_\eta f)\cdot\nabla_\xi-f\nabla_\xi\nabla_\eta+\nabla_{f[\xi,\eta]-(\partial_\eta f)\xi}\\&= fR(\xi,\eta).\end{aligned}$$

类似地成立 $R(\xi,f\eta)=fR(\xi,\eta)$. 最后

$$\begin{aligned}R(\xi,\eta)(f\zeta) &= \nabla_\eta[(\partial_\xi f)\zeta+f\nabla_\xi\zeta]-\nabla_\xi[(\partial_\eta f)\zeta+f\nabla_\eta\zeta]+(\partial_{[\xi,\eta]}f)\zeta+f\nabla_{[\xi,\eta]}\zeta\\&= (\partial_\eta\partial_\xi f)\zeta+(\partial_\xi f)\nabla_\eta\zeta+(\partial_\eta f)\nabla_\xi\zeta+f\nabla_\eta\nabla_\xi\zeta-\\&\quad(\partial_\xi\partial_\eta f)\zeta-(\partial_\eta f)\nabla_\xi\zeta-(\partial_\xi f)\nabla_\eta\zeta-f\nabla_\xi\nabla_\eta\zeta+\\&\quad(\partial_\xi\partial_\eta f-\partial_\eta\partial_\xi f)\zeta+f\nabla_{[\xi,\eta]}\zeta\\&= fR(\xi,\eta)\zeta.\end{aligned}$$

现在由于线性性, 只需对基向量场 $\xi=e_k,\eta=e_l,\zeta=e_j$ 去验证等式 (6),(7) 即可, 其中 $\xi^i=\delta^i_k,\eta^i=\delta^i_l,\zeta^i=\delta^i_j$. 然而对于这样一些向量场, 所要求的等式均由张量 T^j_{kl} 和 R^i_{jkl} 的定义得到. 引理得证. □

应用(四元体系).设在 n 维空间中给出了一个区域, 空间的度量为 g_{ij}. 二次型 $g_{ij}\xi^i\xi^j$ 定义在每个点的切向量上, 我们要将其化为常值形式. 这种约化可以依赖于点光滑地进行. 这表明可以局部地选取 n 个线性无关光滑向量场 $\xi_1,\cdots,\xi_n$, 使得它们两两的内积与点无关:

$$\langle\xi_i,\xi_j\rangle=h_{ij},\quad h_{ij}=\text{常数}. \tag{8}$$

例如, 在广义相对论中出于技术上的方便而选取四个向量 (四元组)ξ_0,ξ_1,ξ_2,ξ_3, 其中 ξ_0 和 ξ_1 为迷向,$(\xi_0)^2=(\xi_1)^2=0,\langle\xi_0,\xi_1\rangle=1$, 而向量 ξ_2,ξ_3 为类空的:$(\xi_2)^2=(\xi_3)^2=-1,\langle\xi_2,\xi_3\rangle=0$.

我们考虑向量场的两两换位子 $[\xi_i,\xi_j]$. 它们可以对这同一基 $\xi_1,\cdots,\xi_n$ 展开为

$$[\xi_i,\xi_j]=c^k_{ij}\xi_k, \tag{9}$$

其中 c^k_{ij} 为展示的 (含变量的) 系数. 成立

定理 30.2. 对于与度量 g_{ij} 相容的对称联络, 成立公式

$$\widehat{\Gamma}^i_{jk} = \frac{1}{2}h^{iq}(c_{jqk} + c_{kqj} - c_{qjk}), \tag{10}$$

其中 $c_{ijk} = h_{is}c^s_{jk}$, 而联络 $\widehat{\Gamma}^i_{jk}$ 的分量由等式

$$\nabla_{\xi_k}\xi_j = \widehat{\Gamma}^i_{jk}\xi_i \tag{11}$$

定义.

证明 由联络的对称性知 $T(\xi_i, \xi_j) = 0$, 又由公式 (6) 知

$$\widehat{\Gamma}^k_{ji} - \widehat{\Gamma}^k_{ij} = c^k_{ij}. \tag{12}$$

另外, 我们以 $\widehat{\Gamma}_{k,ij}$ 表示 $h_{ks}\widehat{\Gamma}^s_{ij}$. 由 (11) 及与度量的相容性得出

$$0 = \nabla_{\xi_k}\langle\xi_i, \xi_j\rangle = \widehat{\Gamma}_{j,ik} + \widehat{\Gamma}_{i,jk}.$$

循环置换指标 i, j, k, 得到三个方程的方程组

$$\widehat{\Gamma}_{i,kj} + \widehat{\Gamma}_{k,ij} = 0,$$

$$\widehat{\Gamma}_{k,ji} + \widehat{\Gamma}_{j,ki} = 0,$$

$$\widehat{\Gamma}_{j,ik} + \widehat{\Gamma}_{i,jk} = 0.$$

解此方程组和 (12) 一起得到了对联络的这些系数所想要的公式. 定理证完. □

利用公式 (7), 可以通过函数 c^k_{ij} 和它们的导数来表达曲率张量 $\langle R(\xi_i, \xi_j)\xi_k, \xi_l\rangle$.

2. 曲率张量的对称性. 由度量产生的曲率张量.

曲率张量具有什么性质?

定理 30.3. 1) $R^i_{qkl} = -R^i_{qlk}$.

2) 对于对称联络具有恒等式:

$$R^i_{qkl} + R^i_{klq} + R^i_{lqk} = 0. \tag{13}$$

3) 对于与度量 g_{ik} 相容的联络, 我们引进张量 $R_{iqkl} = g_{ip}R^p_{qkl}$. 张量 R_{iqkl} 对于指标 i, q 为反对称:

$$R_{iqkl} = -R_{qikl}. \tag{14}$$

4) 对对称联络的曲率张量, 如果此联络与度量 g_{ik} 相容, 则具有对称的形式

$$R_{iqkl} = R_{kliq}. \tag{15}$$

证明 1) 中的等式显然成立. 对公式 (13) 的证明, 我们来计算表达式 $[\nabla_k,\nabla_l]e_q+[\nabla_l,\nabla_q]e_k+[\nabla_q,\nabla_k]e_l$, 其中的 e_q 是基向量场, 并利用等式 $-R^i_{qkl}e_i=[\nabla_k,\nabla_l]e_q$. 我们有 $[\nabla_k,\nabla_l]e_q+[\nabla_l,\nabla_q]e_k+[\nabla_q,\nabla_k]e_l=\nabla_k(\nabla_l e_q-\nabla_q e_l)+\nabla_l(\nabla_q e_k-\nabla_k e_q)+\nabla_q(\nabla_k e_l-\nabla_l e_k)=0$ (由联络的对称性, 每一个括弧内均为零).

我们注意到

$$\begin{aligned}\langle[\nabla_k,\nabla_l]\xi,\xi\rangle &= -g_{ij}R^i_{qkl}\xi^q\xi^j+g_{ij}T^p_{kl}\frac{\partial\xi^i}{\partial x^p}\xi^j\\ &= -R_{jqkl}\xi^q\xi^j+g_{ij}T^p_{kl}\frac{\partial\xi^i}{\partial x^p}\xi^j.\end{aligned}$$

因此要证明 3) 中的对称性只需验证对任意向量场 ξ, 成立 $\langle[\nabla_k,\nabla_l]\xi,\xi\rangle=0$ 即可. 因为此联络与度量相容, 我们有

$$\frac{1}{2}\frac{\partial^2}{\partial x^k\partial x^l}\langle\xi,\xi\rangle=\frac{1}{2}\nabla_k\nabla_l\langle\xi,\xi\rangle=\langle\nabla_k\nabla_l\xi,\xi\rangle+\langle\nabla_l\xi,\nabla_k\xi\rangle,$$

$$\frac{1}{2}\frac{\partial^2}{\partial x^l\partial x^k}\langle\xi,\xi\rangle=\frac{1}{2}\nabla_l\nabla_k\langle\xi,\xi\rangle=\langle\nabla_l\nabla_k\xi,\xi\rangle+\langle\nabla_k\xi,\nabla_l\xi\rangle.$$

因为 $\dfrac{\partial^2}{\partial x^k\partial x^l}=\dfrac{\partial^2}{\partial x^l\partial x^k}$, 那么由上面第一个方程减去第二个便得到了 (14).

最后, 对于由度量产生的曲率张量, 我们来推导恒等式 (15). 在图 33 上由上面每个涂了阴影的面组成的表达式之和, 由于 (13) 和 (14) 的对称性, 它等于零. 对于面 q 和 i 的表达式之和中减去对面 k 和 l 的和便得到了 (15). 定理得证. □

3. 例题: 基灵度量下二维和三维空间的曲率张量

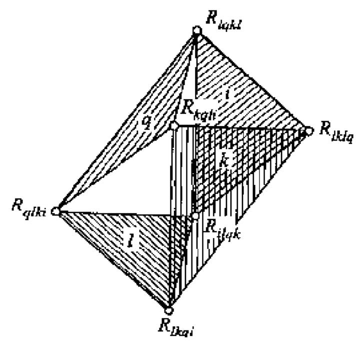


图 33

曲率张量是个 4 阶张量. 作为依赖于反称的一对指标 (k,l) 的对向量场的算子, 得到的自然方式是

$$-R^i_{qkl}T^q=R^i_{qlk}T^q=(\nabla_k\nabla_l-\nabla_l\nabla_k)T^i-T^p_{kl}\frac{\partial T^i}{\partial x^p},$$

其中 $T^p_{kl}=\Gamma^p_{kl}-\Gamma^p_{lk}$ 是挠率张量.

在对称情形有 $T^p_{kl}\equiv 0$. 如果联络为对称并与度量 g_{ij} 相容, 则分量 Γ^k_{ij} 和 R^i_{qkl} 通过 g_{ij} 和它们的导数表示, 又成立下列的对称性:

1) $R^i_{qkl}=-R^i_{qlk}$,

2) $R_{iqkl}=g_{im}R^m_{qkl}=-R_{qikl}$,

3) $R_{iqkl}=R_{kliq}$,

4) $R^i_{qkl}+R^i_{lqk}+R^i_{klq}=0$.

黎曼张量有多少个不同的分量?

在回答这个问题之前我们还要定义由黎曼张量构建的两个张量.

定义 30.1. 称黎曼张量的迹 $R_{ql} = R^i_{qil}$ 为里奇张量.

定义 30.2. 称里奇张量的迹

$$R = g^{lq}R_{ql} = g^{lq}R^i_{qil} \tag{16}$$

为数值曲率.

I. 二维情形. 由对称性 $R_{iqkl} = -R_{iqlk} = -R_{qikl} = R_{kliq}$ 得出结论说, 总存在黎曼张量的一个非零分量, 即 R_{1212}. 所有其余的分量或者由它经指标置换得到, 或者等于零. 有重要的定理:

定理 30.4. 对在三维空间中的二维曲面, 数值曲率 R 与二倍的高斯曲率相等. 故而高斯曲率 (与曲面的平均曲率不同) 可通过这个曲面的黎曼度量表达 (从而是个内蕴不变量).

证明 设曲面由方程 $x = x(u,v), y = y(u,v), z = z(u,v)$ 确定, 其中 x, y, z 为空间的欧氏坐标, 而 $(u,v) = (z^1, z^2)$ 为曲面上的坐标. 在所讨论的点 (非异)$P = (0,0)$ 选取参数 $u = z^1 = x, v = z^2 = y$, 其中 z 轴与曲面垂直. 于是在 P 点附近曲面由方程 $z = f(x,y)$ 描述, 其中 $\mathrm{grad}\, f|_P = 0$. 对于曲面上度量的分量我们有

$$g_{ij} = \delta_{ij} + \frac{\partial f}{\partial z^i}\frac{\partial f}{\partial z^j}, \quad z^1 = x, \quad z^2 = y.$$

特别, 在点 $P = (0,0)$ 处所有的 $\dfrac{\partial g_{ij}}{\partial z^k} = 0$. 所以在此点有 $\Gamma^k_{ij} = 0$ (参看定理 29. 3). 在这样的点上我们有公式 (公式 (2))

$$-R^i_{qkl} = \frac{\partial \Gamma^i_{ql}}{\partial z^k} - \frac{\partial \Gamma^i_{qk}}{\partial z^l},$$

$$R_{iqkl} = \frac{1}{2}\left(\frac{\partial^2 g_{il}}{\partial z^q \partial z^k} + \frac{\partial^2 g_{qk}}{\partial z^i \partial z^l} - \frac{\partial^2 g_{ik}}{\partial z^q \partial z^l} - \frac{\partial^2 g_{ql}}{\partial z^i \partial z^k}\right).$$

由此得出 $(z^1 = x, z^2 = y)$

$$R_{1212} = \frac{1}{2}\left(\frac{\partial^2 g_{12}}{\partial x \partial y} + \frac{\partial^2 g_{12}}{\partial x \partial y} - \frac{\partial^2 g_{11}}{\partial y^2} - \frac{\partial^2 g_{22}}{\partial x^2}\right).$$

此时, $g_{11} = f_x^2 + 1; g_{22} = f_y^2 + 1; g_{12} = f_x f_y$,

$$\left.\frac{\partial^2 g_{11}}{\partial y^2}\right|_P = 2f_{xy}^2, \quad \left.\frac{\partial^2 g_{22}}{\partial x^2}\right|_P = 2f_{xy}^2, \quad \left.\frac{\partial^2 g_{12}}{\partial x \partial y}\right|_P = f_{xx}f_{yy} + f_{xy}^2.$$

最后在点 P 有

$$R_{1212} = f_{xx}f_{yy} - f_{xy}^2 = \det\begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = K.$$

由定义, 在点 P 有 $K = \det\begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$, $\delta_{ij} = g_{ij}$ 在所选的坐标下成立. 但是高斯

曲率在这里等于数值曲率, 而 R_{1212} 为张量的分量. 它们只在所选坐标系下相等, 这时有 $g_{ij}=\delta_{ij}, \det(g_{ij})=1=g$. 容易看出, 由 R 的定义得到 $R=g^{ql}R^i_{qil}$, 我们有

$$R=2\det(g^{ql})R_{1212}=\frac{2}{\det(g_{ij})}R_{1212}=\frac{2}{g}R_{1212}.$$

在我们的坐标系下 $g=1, R_{1212}=K$. 因此在我们坐标系下成立等式 $R=2K$; 因为 R 和 K 为两个数值, 故这等式处处成立. 定理得证. □

注 由此证明可清楚知道, 对黎曼张量的分量有公式

$$R_{1212}=\frac{R}{2}(g_{11}g_{22}-g_{12}^2)=\frac{Rg}{2}=Kg,\quad g=\det(g_{ij}). \tag{17}$$

因此, 高斯曲率 K 为不变量, 而当 $n=2$ 等于 $R/2$, 其中 $R=g^{ql}R^i_{ql}$.

考虑例子.

1) 欧氏度量

$$dl^2=dx^2+dy^2,\quad R^i_{qkl}=0,\quad K=\frac{R}{2}\equiv 0.$$

2) 球面 (见 §9)

$$dl^2=d\rho^2+\sin^2\frac{\rho}{R_0}d\varphi^2;$$

这里 $K=\frac{R}{2}=\frac{1}{R_0^2}>0$ (曲率为正常数).

3) 罗巴切夫斯基平面 (见 §10.1)

$$dl^2=dr^2+\operatorname{sh}^2\frac{r}{R_0}d\varphi^2;$$

这里 $K=\frac{R}{2}=-\frac{1}{R_0^2}<0$ (曲率为负常数).

在 §8.3 中解释过曲率的直观意义, 在那里的曲率是正或负值.

II. 三维情形.这是更为复杂的情形. 黎曼张量

$$R_{iqkl}=-R_{qikl}=-R_{iqlk}=R_{kliq}$$

在这里可以看作在每点的定义在三维线性空间上的二次型, 这个三维空间是由二阶反称张量形成的, 由于对称性, 其维数为 3. 事实上, 如果以 A 表示偶对 $[i,q]=-[q,i]$, 而偶对 $[k,l]=-[l,k]$ 以 B 表示, 则

$$R_{[iq][kl]}=R_{AB}=R_{BA}.$$

因此, 黎曼张量在此由六个数定义:$R_{AB}=R_{BA}, A,B=[1,2],[1,3],[2,3]$. 考虑里奇张量 $R^i_{qil}=R_{ql}=R_{lq}$. 这是个二阶对称张量. 它也由六个数 $R_{ql}, q\geqslant l$ 定义.

数值张量 R 是一个数

$$R = g^{ql}R_{ql} = g^{ql}R^i_{qil}.$$

与二维情形不同, 数值 R 不能完全确定张量 R^i_{qkl}. 但是在三维情形知道了里奇张量就足够了, 理由是我们有公式 (请验证)

$$R_{\alpha\beta\gamma\delta} = R_{\alpha\gamma}g_{\beta\delta} - R_{\alpha\delta}g_{\beta\gamma} + R_{\beta\delta}g_{\alpha\gamma} - R_{\beta\gamma}g_{\alpha\delta} + \frac{R}{2}(g_{\alpha\delta}g_{\beta\gamma} - g_{\alpha\gamma}g_{\beta\delta}). \tag{18}$$

数值曲率是里奇向量的迹 $\mathrm{Tr}(R_{ql}) = g^{ql}R_{ql}$. 我们还有另外的不变量, 即特征值 $\lambda_1, \lambda_2, \lambda_3$, 由下面的方程决定:

$$\det(R_{ql} - \lambda g_{ql}) = 0, \tag{19}$$

而 $\lambda_1 + \lambda_2 + \lambda_3 = R$.

(当说"正曲率空间"时等于说黎曼张量 R_{AB} 为在二阶反称张量上的二次型为正定.)

III. *四维情形*. 这时里奇张量不能确定黎曼. 尽管如此, 里奇张量仍非常重要. 例如在四维时空中假定引力场为度量 $(g_{ij}), i, j = 0, 1, 2, 3$, 而所有其他的物质的性质都集中于"能量-动量张量"λT_{ij} (λ 为表示物质大小的常数).

定义时空空间度量的爱因斯坦方程具形式

$$R_{ij} - \frac{1}{2}Rg_{ij} = \lambda T_{ij}, \quad \nabla_j T^j_i = 0. \tag{20}$$

在物质缺失的情形下有

$$R_{ij} - \frac{1}{2}Rg_{ij} = 0 \ (\text{或者 } R_{ij} = 0). \tag{21}$$

在此时 $\det(g_{ij}) \neq 0$, 然而度量为不定的 (它在对角形中三个为负, 一个为正; 见 §37).

IV. *基灵度量的曲率张量*. 设 G 为矩阵变换群, $\mathfrak{g}$ 为其李代数, 又在 $\mathfrak{g}$ 上给出了基灵度量. 在群 G 上令

$$\nabla_{L_X}L_Y = \frac{1}{2}L_{[X,Y]} = \frac{1}{2}[L_X, L_Y], \tag{22}$$

以引进联络, 其中 $X, Y \in \mathfrak{g}$, L_X, L_Y 为相应的左不变向量场, 即 $L_X(A) = AX$, $L_Y(A) = AY$, 其中 $A \in G$ 为任意元 (见 §24.4).

形如 L_X 的场, $X \in \mathfrak{g}$, 构成 G 在每点切空间的基 (参看引理 24.3), 所以公式 (22) 完全决定了一个联络 (假设满足莱布尼茨公式; 参看 §28).

引理 30.2. 联络 (22) 为对称的, 并且与群 G 上基灵度量相容.

证明 验证对于李代数 $\mathfrak{g}$ 中的任意元 X, Y 有 $T(L_X, L_Y) = 0$. 事实上 (公式 (6)),

$$\begin{aligned} T(L_X, L_Y) &= \nabla_{L_X}L_Y - \nabla_{L_Y}L_X - [L_X, L_Y] \\ &= \frac{1}{2}L_{[X,Y]} - \frac{1}{2}L_{[Y,X]} - L_{[X,Y]} = 0. \end{aligned}$$

我们现在证明它与基灵度量 $\langle,\rangle$ 的相容性. 只要证明对任意向量场 ξ,η,ζ 有下面关于内积的微分规则即可:

$$\partial_\xi\langle\eta,\zeta\rangle=\langle\nabla_\xi\eta,\zeta\rangle+\langle\eta,\nabla_\xi\zeta\rangle.$$

这只要对 ξ,η,ζ 为左不变向量场去验证这个等式.这时我们有$\langle L_Y(A),L_Z(A)\rangle=\langle Y,Z\rangle_0=$ 常数, 即不依赖于 $A\in G$, 因而

$$\partial_{L_X}\langle L_Y,L_Z\rangle\equiv 0.$$

(这里的 $\langle,\rangle_0$ 是相应的李代数 $\mathfrak{g}$ 上基灵度量.) 另一方面,

$$\begin{aligned}\langle\nabla_{L_X}L_Y,L_Z\rangle+\langle L_Y,\nabla_{L_X}L_Z\rangle &=\frac{1}{2}\{\langle L_{[X,Y]},L_Z\rangle+\langle L_Y,L_{[X,Z]}\rangle\}\\ &=\frac{1}{2}\{\langle[X,Y],Z\rangle_0+\langle Y,[X,Z]\rangle_0\}=0,\end{aligned}$$

这是因为 ad X 关于基灵度量是个反称的线性算子 (见 §24). 引理得证. □

由此 (以及由雅可比恒等式) 容易得到

推论 与基灵度量相容的对称联络的曲率有公式 (参照公式 (7)) 为

$$R(L_X,L_Y)L_Z=\frac{1}{4}L_{[[X,Y],Z]},\quad \langle R(L_X,L_Y)L_Z,L_W\rangle=\frac{1}{4}\langle[X,Y],[Z,W]\rangle_0. \tag{23}$$

我们还要求出与基灵度量相容的联络的测地线. 因为群 G 上的位移是平移, 故而只要确定通过群的单位元的测地线就可以了. 有下面的定理.

定理 30.5. 通过群单位元的基灵度量的测地线正好是群的单参数子群.

证明 设 $A(t)=\exp(tX),X\in\mathfrak{g},t$ 为实参数, 为单参数子群. 它的速度向量是左不变向量场 L_X(更准确地说, 它是在曲线 $A(t)$ 上的限制). 故而

$$\nabla_{\dot A}\dot A=\nabla_{L_X}L_X=\frac{1}{2}L_{[X,X]}=0 \tag{24}$$

因此, 此时的单参数子群是测地线. 因为从单位元可以发出具任意初始速度向量的单参数子群, 故而由唯一性定理我们得到了所有的测地线. 定理证毕. □

4. 彼得松 - 柯达齐方程. 具常负曲率的曲面和 "正弦–戈登" 方程

设 $r=r(x^1,x^2)$ 为在欧氏空间 $\mathbb{R}^3$ 中的曲面,g_{ij} 为在曲面上诱导的度量. 在曲面上也出现了第二基本型 $b_{ij}dx^idx^j$, 其中

$$b_{ij}=\left\langle\frac{\partial}{\partial x^j}\left(\frac{\partial r}{\partial x^i}\right),n\right\rangle, \tag{25}$$

n 为曲面的单位法向量,$\langle,\rangle$ 为 $\mathbb{R}^3$ 中的欧几里得内积.

如何计算与度量 g_{ij} 相容的对称联络的分量?

命题 30.1. 设在曲面 $r(x^1,x^2)$ 上的度量为 g_{ij}, 那么与此度量相容的对称联络的分量可用下面公式计算:

$$\Gamma_{ij}^k=\left\langle\frac{\partial^2 r}{\partial x^i\partial x^j},\frac{\partial r}{\partial x^l}\right\rangle g^{lk}\quad(i,j,k=1,2),\tag{26}$$

或者

$$\frac{\partial^2 r}{\partial x^i\partial x^j}=b_{ij}n+\Gamma_{ij}^k\frac{\partial r}{\partial x^k}.\tag{27}$$

证明 因为 $g_{ij}=\left\langle\frac{\partial r}{\partial x^i},\frac{\partial r}{\partial x^j}\right\rangle$ (参看 §7.3), 则

$$\frac{\partial g_{ij}}{\partial x^s}=\left\langle\frac{\partial^2 r}{\partial x^i\partial x^s},\frac{\partial r}{\partial x^j}\right\rangle+\left\langle\frac{\partial r}{\partial x^i},\frac{\partial^2 r}{\partial x^j\partial x^s}\right\rangle.$$

由这个等式和克里斯托费尔公式 (29.3.13), 对 $i,j,s=1,2$ 得到 (26).

$\frac{\partial^2 r}{\partial x^i\partial x^j}$ 沿方向 n 的分量为 b_{ij}, 另外, 向量 n 正交于 $\frac{\partial r}{\partial x^i}$ 和 $\frac{\partial r}{\partial x^j}$, 故而对某个 x_{ij}^k 有

$$\frac{\partial^2 r}{\partial x^i\partial x^j}=b_{ij}n+x_{ij}^k\frac{\partial r}{\partial x^k},$$

又,

$$\left\langle\frac{\partial^2 r}{\partial x^i\partial x^j},\frac{\partial r}{\partial x^l}\right\rangle=x_{ij}^k\left\langle\frac{\partial r}{\partial x^k},\frac{\partial r}{\partial x^l}\right\rangle=x_{ij}^k g_{kl}.$$

比较这最后的等式与 (26), 得到 $x_{ij}^k=\Gamma_{ij}^k$. 这就是所要证明的. □

简明起见, 记 $\frac{\partial r}{\partial x^i}=e_i, i=1,2$. 向量 (e_1,e_2,n) 构成了 $\mathbb{R}^3$ 中的正交基, 它们光滑地依赖于曲面的点. 我们还要算出导数 $\frac{\partial n}{\partial x^i}, i=1,2$. 因为向量 n 为单位法向量, 故导数 $\frac{\partial n}{\partial x^i}$ 与 n 正交 (参看 §5). 由等式 $\langle n,e_j\rangle=0$ 得到

$$\begin{aligned}0=\frac{\partial}{\partial x^i}\langle n,e_j\rangle&=\left\langle\frac{\partial n}{\partial x^i},e_j\right\rangle+\left\langle n,\frac{\partial e_j}{\partial x^i}\right\rangle\\&=\left\langle\frac{\partial n}{\partial x^i},e_j\right\rangle+b_{ij},\end{aligned}$$

由此得出

$$\frac{\partial n}{\partial x^i}=-b_i^j e_j,\quad b_i^j=g^{jl}b_{il}.\tag{28}$$

由等式 (27) 和 (28) 得到下面的相容性的条件:

$$\begin{aligned}\frac{\partial^2 e_i}{\partial x^k \partial x^j} &= \frac{\partial b_{ij}}{\partial x^k} n - b_{ij} b_k^l e_l + \frac{\partial \Gamma_{ij}^l}{\partial x^k} e_l + \Gamma_{ij}^s [b_{ks} n + \Gamma_{ks}^l e_l] \\ &= \frac{\partial b_{ik}}{\partial x^j} n - b_{ik} b_j^l e_l + \frac{\partial \Gamma_{ik}^l}{\partial x^j} e_l + \Gamma_{ik}^s [b_{js} n + \Gamma_{js}^l e_l] \\ &= \frac{\partial^2 e_i}{\partial x^j \partial x^k},\end{aligned}$$

或者

$$\frac{\partial \Gamma_{ij}^l}{\partial x^k} - \frac{\partial \Gamma_{ik}^l}{\partial x^j} + \Gamma_{ij}^s \Gamma_{ks}^l - \Gamma_{ik}^s \Gamma_{js}^l = b_{ij} b_k^l - b_{ik} b_j^l, \tag{29}$$

$$\frac{\partial b_{ij}}{\partial x^k} - \frac{\partial b_{ik}}{\partial x^j} = \Gamma_{ik}^s b_{js} - \Gamma_{ij}^s b_{ks}. \tag{30}$$

称方程 (29) 为*高斯方程*.(29) 的左端即是曲率张量 R_{ijk}^l, 并且这个等式等价于上面证明过的关于高斯曲率和数值曲率之间关系的定理 (定理 30.4; 请验证!)

称公式 (30) 为*彼得松 – 柯达齐方程*.

注 彼得松 – 柯达齐方程给出了使形式 $b_{ij}(x^1, x^2)$ 能够成为空间 $\mathbb{R}^3$ 中曲面上的第二基本型的必要条件, 而曲面上的度量为 $g_{ij}(x^1, x^2)$, 在这里的 Γ_{ij}^k 由克里斯托费尔公式用 g_{ij} 计算. 可以证明这个条件也是充分的.

设一个曲面具负曲率 $K < 0$, 从而 $b_{12}^2 - b_{11} b_{22} > 0$ (参看 §8.3 中的公式 (26)). 于是能够在此曲面上 (局部地) 引入这样的坐标 (p, q), 使得在此坐标下第二基本型为

$$b_{ij} dx^i dx^j = 2 b_{pq} dp dq \tag{31}$$

(缺少了 dp^2 和 dq^2 的项). 如果另外有 $K=$ 常数 (例如 $K = -1$), 则对方程 $K(g_{pp} g_{qq} - g_{pq}^2) = -b_{pq}^2$ 微分并利用彼得松 – 柯达齐方程, 我们得到 (请验证!)

$$\frac{\partial g_{pp}}{\partial q} = 0, \quad \frac{\partial g_{qq}}{\partial p} = 0 \tag{32}$$

(在这里注意到了在坐标 p, q 的选取中可以任意, 只要使得局部有 $g_{pq} \neq 0$ 就可以了). 我们选取新的坐标 (x, y) 于曲面上, 其中

$$x = \int_{p_0}^{p} \sqrt{g_{pp}} dp, \quad y = \int_{q_0}^{q} \sqrt{g_{qq}} dq. \tag{33}$$

在坐标 (x, y) 下第一和第二基本型为

$$ds^2 = dx^2 + 2 g_{xy} dx dy + dy^2, \quad b_{ij} dx^i dx^j = 2 b_{xy} dx dy. \tag{34}$$

设 $g_{xy} = \cos \omega$, 其中 ω 为坐标曲线 (x, y) 之间的夹角 (切线的夹角). 于是在 $K \equiv -1$ 时高斯方程为

$$\omega_{xy} = \sin \omega. \tag{35}$$

在物理学的文献中常称此方程为“正弦 – 戈登”方程. 令 $x=\frac{\tau+\xi}{\sqrt{2}}, y=\frac{\tau-\xi}{\sqrt{2}}$, 此方程化为

$$\frac{\partial^2\omega}{\partial\tau^2}-\frac{\partial^2\omega}{\partial\xi^2}=\sin\ \omega. \tag{36}$$

习题

30.1. 证明方程 (36) 不依赖于 ξ, 且当 $\tau\to+\infty$ 时递降的解对应于曲率 $K=-1$ 的如下旋转面:

$$z=-\sqrt{1-x^2-y^2}+\ln\frac{1+\sqrt{1-x^2-y^2}}{\sqrt{x^2+y^2}}$$

(贝尔特拉米伪球).

30.2. 证明公式 (18).

30.3. 设 $x^i(t), i=1,2$ 为分段光滑曲线, 它围成了区域 U. 证明 $\Delta\varphi=\iint\limits_U K\sqrt{g}dx^1\wedge dx^2$ 为向量沿曲线 $x^i(t)$ 平行移动一周的旋转角 (K 为高斯曲率).

30.4. 如果这条曲线由三条测地线弧组成, 并且曲率 K 为常数, 则这个测地三角形的角之和等于 $\pi+K\sigma$, 其中 σ 为此三角形的面积 (请证明!). 再考察球面和罗巴切夫斯基平面的情形.

30.5. 设 $\xi_1,\cdots,\xi_n$ 是在 n 维黎曼 (或伪黎曼) 空间中的向量场满足 $g_{ij}=\langle\xi_i,\xi_j\rangle, [\xi_i,\xi_j]=c_{ij}^k\xi_k$. 计算与此度量相容的对称联络 Γ_{ij}^k(这里 $\nabla_{\xi_j}\xi_i=\Gamma_{ij}^k\xi_k$).

30.6. 我们将向量 $\xi=(\xi^k)$ 绕正方形平移一周; 此正方形的边为 ε, 由坐标轴 x^i, x^j 张成 (逆时针方向). 设 $\widetilde{\xi}(\varepsilon)$ 为此绕行一圈后的结果. 证明

$$\lim_{\varepsilon\to0}\frac{\widetilde{\xi}^k(\varepsilon)-\xi^k}{\varepsilon^2}=-R_{lij}^k\xi^l.$$

30.7. 证明对与度量相容的对称联络的曲率张量的比安基恒等式

$$\nabla_mR_{ikl}^n+\nabla_lR_{imk}^n+\nabla_kR_{ilm}^n=0.$$

30.8. 由上个习题中的公式推出下面对里奇张量的发散量的恒等式:

$$\nabla_lR_m^l=\frac{1}{2}\frac{\partial R}{\partial x^m}.$$

30.9. 设 $X_1,\cdots,X_n$ 为 n 维黎曼空间中的法正交向量场, $\omega_1,\cdots,\omega_n$ 为对偶 1– 形式: $\omega_i(X_j)=\delta_{ij}$ (假定所有指标均为下指标). 定义 1– 形式 ω_{ij} 和 2– 形式 Ω_{ij} 为

$$\omega_{ij}=\Gamma_{jk}^i\omega_k;\quad \Omega_{ij}=\frac{1}{2}R_{ijkl}\omega_k\wedge\omega_l.$$

其中 $\nabla_{X_k}X_j = \Gamma^i_{jk}X_i, \langle R(X_k, X_l)X_j, X_i\rangle = R_{ijkl}$; 而取和是对重复指标进行的.

a) 证明 $\omega_{ij} = -\omega_{ji}$,

b) 推导下面的关系 (嘉当结构方程)

$$d\omega_i = -\omega_j \wedge \omega_{ij},$$

$$d\omega_{ij} = \omega_{il} \wedge \omega_{lj} - \Omega_{ij},$$

$$d\Omega_{ij} = -\Omega_{il} \wedge \omega_{lj} + \omega_{il} \wedge \Omega_{lj}.$$

30.10. 在上一习题的记号下, 我们定义在偶维数 n 情形下的形式 $\Omega_{(k)}$ 和 Ω 如下:

$$\Omega_{(k)} = \Omega_{i_1 i_2} \wedge \Omega_{i_2 i_3} \wedge \cdots \wedge \Omega_{i_{k-1} i_k} \wedge \Omega_{i_k i_1},$$

$$\Omega = \varepsilon^{i_1 \cdots i_n} \Omega_{i_1 i_2} \wedge \Omega_{i_3 i_4} \wedge \cdots \wedge \Omega_{i_{n-1} i_n} \ (n = 2m).$$

a) 证明形式 $\Omega_{(k)}, \Omega$ 的定义不依赖于法正交标架 $X_1, \cdots, X_n$.

b) 形式 $\Omega_{(k)}, \Omega$ 为闭.

c) 得出这些形式在坐标下的表达式.

d) 对 $n = 2$, 形式 Ω 为 $\Omega = K\sqrt{g}dx^1 \wedge dx^2, K$ 为高斯曲率.

e) 对伪黎曼空间推导出类似于习题 30.9 和 30.10 的公式.

第五章　变分法原理

§31. 一维变分问题

1. 欧拉 – 拉格朗日方程

我们在 §29 中定义过测地线 $x^i = x^i(t)$, 其方程为 $\nabla_T(T) = 0$, 其中 $T^i = \dfrac{dx^i}{dt}$ 为曲线的速度向量, 或者

$$\frac{d^2x^i}{dt^2} + \Gamma^i_{jk}\frac{dx^j}{dt}\frac{dx^k}{dt} = 0. \tag{1}$$

如果联络 Γ^i_{jk} 对称且与度量 g_{ij} 相容, 则分量 Γ^i_{jk} 可通过 g_{ij} 表达, 从而测地线由度量定义. 除了测地线的速度向量沿自身平移仍为速度向量外, 测地线还有其他什么几何性质? 我们所熟知的是测地线 (局部地) 最短, 即它比连接两个相互充分靠近同样点的任意其他曲线的长度更小. 我们要在这里解释这个问题.

以更广泛的观点对待这个问题是有好处的. 设 $L(x, \xi, t)$ 为点 $x = (x^1, \cdots, x^n)$ 的任意函数, 而 $\xi = (\xi^i)$ 为在此点的切向量. 我们考虑一对固定的点 $P = (x_1^1, \cdots, x_1^n)$ 和 $Q = (x_2^1, \cdots, x_2^n)$, 和所有连接这两个点的光滑曲线 $\gamma : x^i = x^i(t), a \leqslant t \leqslant b$ (固定的 a 和 b), $x^i(a) = x_1^i, x^i(b) = x_2^i$.

考虑量

$$S[\gamma] = \int_P^Q L(x(t), \dot{x}(t), t)dt. \tag{2}$$

量 (泛函)$S[\gamma]$ 被称做作用. 在哪条曲线 γ 上泛函 $S[\gamma]$ 极小?

例 31.1. 设 $L(x, \xi) = g_{ij}\xi^i\xi^j$. 于是 $S[\gamma] = \displaystyle\int_P^Q L(x, \dot{x})dt = \int_P^Q g_{ij}\dot{x}^i\dot{x}^j dt =$

$\int_P^Q |\dot{x}|^2 dt$. 在哪条曲线 $\gamma = \{x(t)\}$ 上泛函 $S[\gamma]$ 极小?

例 31.2. 设 $L(x,\xi) = \sqrt{g_{ij}\xi^i\xi^j} = \sqrt{\langle\xi,\xi\rangle} = |\xi|$(向量的长度). 于是 $S[\gamma] = \int_P^Q \sqrt{g_{ij}\dot{x}^i\dot{x}^j}dt$ 为曲线 γ 的长. 在点 P 和 Q 间哪一条曲线长度最小?

例 31.3. 设度量是欧几里得的. 令 $L = \frac{m}{2}\delta_{ij}\xi^i\xi^j - U(x)$, 其中 U 为点的某个函数. 于是 $S[\gamma] = \int_P^Q [\sum_i \frac{m}{2}(\dot{x}^i)^2 - U(x)]dt$. 使 $S[\gamma]$ 极小的曲线 γ 是质量为 m 的在力场 $f_i = -\frac{\partial U}{\partial x^i}$ 中的运动轨线.

我们有一个简明的定理.

定理 31.1. 如果 $S[\gamma] = \int_P^Q L(x,\dot{x},t)dt$ 在从 P 出发到 Q 的所有光滑曲线中的曲线 $\gamma : x^i = x^i(t)$ 上取得极小值, 则沿曲线 γ 成立方程

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) - \frac{\partial L}{\partial x^i} = 0, \quad i = 1,\cdots,n, \tag{3}$$

其中

$$\frac{\partial L}{\partial \dot{x}^i} = \left.\frac{\partial L(x,\xi,t)}{\partial \xi^i}\right|_{\xi=\dot{x}},$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) = \left.\left(\frac{\partial^2 L}{\partial\xi^i\partial\xi^j}\ddot{x}^j + \frac{\partial^2 L}{\partial\xi^i\partial x^j}\dot{x}^j + \frac{\partial^2 L}{\partial\xi^i\partial t}\right)\right|_{\xi=\dot{x}}$$

(看作是 $L = L(x^1,\cdots,x^n,\xi^1,\cdots,\xi^n,t)$, 其中 x 和 ξ 为独立变量, 然后再在曲线上以 $\xi^i = \frac{dx^i}{dt}$ 代入).

证明 设 $\eta^i = \eta^i(t), a\leqslant t\leqslant b$, 为任意光滑函数使得 $\eta^i(a) = 0, \eta^i(b) = 0$. 考虑表达式

$$\lim_{\varepsilon\to 0}\frac{S[\gamma+\varepsilon\eta] - S[\gamma]}{\varepsilon} = \left.\frac{d}{d\varepsilon}S[\gamma+\varepsilon\eta]\right|_{\varepsilon=0}.$$

其中 $\gamma+\varepsilon\eta$ 为曲线 $x^i = x^i(t) + \varepsilon\eta^i(t)$, 它也由 P 到 Q 并以小的 ε 靠近曲线 $\gamma(t)$.

引理 31.1. 如果 $S[\gamma]$ 极小, 对任意光滑的向量值函数 $\eta(t)$, 如果它在时间区间的两端取零值, 我们则有

$$\lim_{\varepsilon\to 0}\frac{S[\gamma+\varepsilon\eta] - S[\gamma]}{\varepsilon} = \left.\frac{d}{d\varepsilon}S[\gamma+\varepsilon\eta]\right|_{\varepsilon=0} \equiv 0.$$

证明 显然. □

再回到定理的证明. 我们来展开 $\frac{d}{d\varepsilon}S[\gamma+\varepsilon\eta]|_{\varepsilon=0}$ 我们有

$$\frac{d}{d\varepsilon}S[\gamma+\varepsilon\eta]\bigg|_{\varepsilon=0}=\int_a^b\left\{\frac{\partial L}{\partial x^i}\eta^i(t)+\frac{\partial L}{\partial \xi^i}\dot{\eta}^i(t)\right\}dt=0, \tag{4}$$

其中的积分, 按定义应沿曲线 γ:

$$x^i=x^i(t),\quad \xi^i=\dot{x}^i(t)$$

进行计算.

这个等式对于在时间区间两端为零的任意光滑向量值函数都成立.

我们注意有下面的恒等式:

$$\int_a^b\frac{\partial L}{\partial \xi^i}\dot{\eta}^i dt=\frac{\partial L}{\partial \xi^i}\eta^i\bigg|_{t=b}-\frac{\partial L}{\partial \xi^i}\eta^i\bigg|_{t=a}-\int_a^b\eta^i\frac{d}{dt}\left(\frac{\partial L}{\partial \xi^i}\right)dt$$

(分部积分). 因为 $\eta^i(a)=\eta^i(b)=0$, 我们得到 $\displaystyle\int_a^b\frac{\partial L}{\partial \xi^i}\dot{\eta}^i dt=-\int_a^b\frac{d}{dt}\left(\frac{\partial L}{\partial \xi^i}\right)\eta^i dt.$ 将此表达式代入公式 (4), 我们看出, 对任意在时间区间两端为零的光滑向量值函数 $\eta^i(t)$, 成立等式

$$\frac{d}{d\varepsilon}S[\gamma+\varepsilon\eta]\bigg|_{\varepsilon=0}=\int_a^b\left[\frac{\partial L}{\partial x^i}-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}^i}\right]\eta^i dt=0, \tag{5}$$

这时, 我们记得在所有联结点 P 和 Q 的光滑曲线中, 它在曲线 $\gamma:x^i=x^i(t)$ 达到函数 $S[\gamma]$ 的极小值.

由此可以得出等式

$$\psi^i(t)=\frac{\partial L}{\partial x^i}-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}^i}=0,\quad i=1,\cdots,n.$$

事实上, 如果 $\psi^i(t)\neq 0$ 对某个 i 和某个 a 与 b 之间的 t_0 成立, 则容易选出那样的函数 $\eta^i(t)$ 使 (5) 不为零 (例如, 令 $\eta^i=\psi^i f(t)$, 当 $f(t)\geqslant 0$ 并在区间端点取零时我们将在积分下得到正数). 因而定理得证. □

定义 31.1. 称方程 $\dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{x}^i}\right)=\dfrac{\partial L}{\partial x^i}$ 的解为泛函 S 的极值曲线.

我们再给出几个定义.

1) 称被积函数

$$L=L(x,\xi,t)=L(x,\dot{x},t) \tag{6}$$

为拉格朗日函数.

2) 称表达式

$$E=E(x,\dot{x},t)=E(x,\xi,t)=\xi^i\frac{\partial L}{\partial \xi^i}-L=\dot{x}^i\frac{\partial L}{\partial \dot{x}^i}-L \tag{7}$$

为能量.

3) 称表达式

$$p_i = \frac{\partial L}{\partial \dot{x}^i} = \frac{\partial L}{\partial \xi^i}\ (\text{余向量}) \tag{8}$$

为动量.

4) 称表达式

$$f_i = \frac{\partial L}{\partial x^i}\ (\text{也是余向量}) \tag{9}$$

为力.

5) 称定理 31.1 中的方程 (极值方程)

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) = \frac{\partial L}{\partial x^i} \quad \text{或} \quad \dot{p}_i = f_i$$

为欧拉—拉格朗日方程.

6) 称表达式

$$\frac{\delta S}{\delta x^i} = \frac{\partial L}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) \tag{10}$$

为泛函 $S[\gamma]$ 的变分导数.由定理 31.1 的证明中得出另一个关于变分导数的定义: 量 $\frac{\delta S}{\delta x^i}$ 由下面等式定义:

$$\left.\frac{d}{d\varepsilon}S[\gamma+\varepsilon\eta]\right|_{\varepsilon=0} = \int_a^b \frac{\delta S}{\delta x^i}\eta^i dt. \tag{11}$$

应该指出, 拉格朗日函数, 能量和动量的定义并非完全确定的, 它只准确到一种变换形式:

$$L' = L + \frac{df(x,t)}{dt}, \quad E' = E - \frac{\partial f}{\partial t}, \quad p' = p + \frac{\partial f}{\partial x}.$$

2. 泛函的基本例子

例 31.4. 如果 $L = \frac{m}{2}\sum_i (\xi^i)^2 - U(x)$, 其中 U 为点的函数, 则 $f_i = -\frac{\partial U}{\partial x^i}, p_i = m\dot{x}^i$. 我们有

$$\dot{p}_i = f_i \quad \text{或} \quad m\ddot{x}^i = -\frac{\partial U}{\partial x^i}. \tag{12}$$

这是在位势力场 $f = -\operatorname{grad} U$ 下, 质量为 m 的质点的运动 (牛顿) 方程. 我们在经典力学中已熟知了. 我们得到这个质点的运动轨线与泛函 S 的极值曲线相同. 其中

$$S = \int \left[\frac{m}{2}\sum(\dot{x}^i)^2 - U(x)\right] dx$$

(最小作用原理).

例 31.5. 如果 $L = \frac{1}{2}g_{ij}\xi^i\xi^j$, 则 $p_i = g_{ij}\xi^j, f_k = \frac{1}{2}\frac{\partial g_{ij}}{\partial x^k}\xi^i\xi^j$. 我们得到极值方程

$$\frac{dp_k}{dt} = \frac{1}{2}\frac{\partial g_{ij}}{\partial x^k}\dot{x}^i\dot{x}^j, \quad \text{其中} \quad p_k = g_{kj}\dot{x}^j,$$

即
$$\frac{dp_k}{dt}=\ddot{x}^j g_{jk}+\dot{x}^j\frac{\partial g_{jk}}{\partial x^i}\dot{x}^i=\frac{1}{2}\frac{\partial g_{ij}}{\partial x^k}\dot{x}^i\dot{x}^j.$$
因为 $g^{km}g_{jk}=\delta_j^m$, 我们得到
$$\ddot{x}^m+g^{km}\left(\frac{\partial g_{jk}}{\partial x^i}-\frac{1}{2}\frac{\partial g_{ij}}{\partial x^k}\right)\dot{x}^i\dot{x}^j=0.$$
现在注意到下面的恒等式:
$$g^{km}\frac{\partial g_{jk}}{\partial x^i}\dot{x}^i\dot{x}^j=\frac{1}{2}\dot{x}^i\dot{x}^j g^{km}\left(\frac{\partial g_{ik}}{\partial x^j}+\frac{\partial g_{jk}}{\partial x^i}\right).$$
将其代入上面的方程中我们便得到了
$$\ddot{x}^m+\Gamma_{ij}^m\dot{x}^i\dot{x}^j=0, \tag{13}$$
其中
$$\Gamma_{ij}^m=\frac{1}{2}g^{km}\left(\frac{\partial g_{ik}}{\partial x^j}+\frac{\partial g_{jk}}{\partial x^i}-\frac{\partial g_{ij}}{\partial x^k}\right). \tag{14}$$
这是与度量 g_{ij} 相容的对称联络.

因此, 我们已证明了

定理 31.2. 如果 $L=g_{ij}\dot{x}_i\dot{x}_j=\langle\dot{x},\dot{x}\rangle=|\dot{x}|^2, S[\gamma]=\int_P^Q[\dot{x}]^2dt$, 则极值曲线 (特别地, 极小) 的欧拉 – 拉格朗日方程与测地线方程相同.

例 31.6. 如果 $L=\sqrt{g_{ij}\dot{x}^i\dot{x}^j}=|\dot{x}|$, 则表达式 $S=\int_P^Q|\dot{x}|dt$(长度) 不依赖于参数 t. 这时的欧拉 – 拉格朗日方程为 $\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{x}^k}\right)=\frac{\partial L}{\partial x^k}$, 或者
$$\frac{d}{dt}\left(\frac{g_{kj}\dot{x}^j}{\sqrt{g_{ij}\dot{x}^i\dot{x}^j}}\right)=\frac{\dfrac{\partial g_{ij}}{\partial x^k}\dot{x}^i\dot{x}^j}{2\sqrt{g_{ij}\dot{x}^i\dot{x}^j}}.$$
如果赋予曲线的参数与自然参数成比例, 即 $t=$ 常数 $\cdot l$, 则 $\sqrt{g_{ij}\dot{x}^i\dot{x}^j}=|\dot{x}|=$ 常数, 于是得到
$$\frac{d}{dl}(g_{kj}\dot{x}^j)=\frac{1}{2}\frac{\partial g_{ij}}{\partial x^k}\dot{x}^i\dot{x}^j,$$
这与上面例 31.5 的极值方程相同, 但现在只在曲线的参数与自然参数成比例时得到.

在这章我们将约定以更广的含义来解释 “自然参数”这个词汇:我们也称与自然参数成比例的参数为自然参数.

我们注意, 在例 31.2 中也可局限于考虑具自然参数的曲线, 这是由于曲线的长度不依赖于曲线的参数.

那么, 我们已证明了

定理 31.3. 如果在曲线上选取自然参数, 则对长度 $L=\sqrt{g_{ij}\xi^i\xi^j}$ 的泛函的极值 (特别是极小) 曲线的欧拉 – 拉格朗日方程同于测地线的方程. 因此作为连接点 P 和 Q 的所有曲线中最短的光滑曲线满足对于自然参数的测地线方程.

我们来指出任意拉格朗日函数的能量和动量的某些性质.

第一个性质 ("能量守恒"). 如果拉格朗日函数 $L=L(x,\xi)$ 与 t 无显式关系, 则能量 E 沿极值曲线的全导数等于零:

$$\frac{dE}{dt}=\frac{d}{dt}\left(\dot{x}^i\frac{\partial L}{\partial \dot{x}^i}-L\right)=\ddot{x}^i\frac{\partial L}{\partial \dot{x}^i}+\dot{x}^i\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right)-\frac{\partial L}{\partial x^i}\dot{x}^i-\frac{\partial L}{\partial \dot{x}^i}\ddot{x}^i$$
$$=\dot{x}^i\left(\frac{d}{dt}\frac{\partial L}{\partial \dot{x}^i}-\frac{\partial L}{\partial x^i}\right)\equiv 0.$$

第二个性质 ("动量守恒"). 如果选择坐标 $x^1,\cdots,x^n$ 使 $\dfrac{\partial L}{\partial x^i}\equiv 0$, 则沿任意极值曲线成立等式 $\dot{p}_i=\dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{x}^i}\right)\equiv 0$. 这可由欧拉 – 拉格朗日方程得到. 在这种情况称坐标 x^i 为循环坐标.

例 31.7 如果 $L=\dfrac{1}{2}g_{ij}\dot{x}^i\dot{x}^j$, 于是 $E=L=\dfrac{1}{2}|\dot{x}|^2$. 由能量守恒定律, 沿泛函 $S=\int L dt$ 的极值曲线有 $\dfrac{dE}{dt}=0$. 因此, 这时极值曲线总是测地线, 而沿整个曲线其速度都为常数 (在自然参数下的速度).

注 如果拉格朗日函数 $L(x,\dot{x})$ 是关于 $\xi=\dot{x}$ 的一阶齐次函数, 即 $L(x,\lambda\xi)=\lambda L(x,\xi)$ (例如, $L=\sqrt{\langle\xi,\xi\rangle}$), 则能量 E 恒等于零, 且在极值曲线上可取任意参数.

例 31.8 如果三维欧氏空间中的曲面由柱面坐标的方程 $f(z,r)=0$ 给出 (旋转面), 则在曲面上 (局部) 坐标中的一个可取作角 φ, 而另一个为 r 或 z. 度量的形式为

$$dl^2=dz^2+dr^2+r^2d\varphi^2.$$

若局部地有 $r=r(z)$, 则

$$dl^2=g_{zz}dz^2+r^2(z)d\varphi^2,\quad g_{ij}=\begin{pmatrix} g_{zz} & 0\\ 0 & r^2\end{pmatrix}.$$

对于测地线的拉格朗日函数为

$$L=\frac{1}{2}(g_{zz}\dot{z}^2+r^2(z)\dot{\varphi}^2)=\frac{1}{2}(g_{rr}\dot{r}^2+r^2\dot{\varphi}^2),$$

又有能量 $E=L$ 和动量 $p_\varphi=\dfrac{\partial L}{\partial\dot{\varphi}}=r^2\dot{\varphi},p_z=\dfrac{\partial L}{\partial \dot{z}}=g_{zz}\dot{z}$ 守恒.

在曲面上的局部坐标为 z 和 φ. 考虑它们的单位基向量 e_z 和 e_φ. 这些基向量的内积为:

$$\langle e_z, e_z\rangle = g_{zz}, \quad \langle e_z, e_\varphi\rangle = 0, \quad \langle e_\varphi, e_\varphi\rangle = r^2(z).$$

考虑测地线的速度向量 $v = (\dot{z}, \dot{\varphi})$, 我们来计算 v 和 e_φ 之间的夹角 ψ:

$$\cos\ \psi = \frac{\langle v, e_\varphi\rangle}{\sqrt{\langle v, v\rangle\langle e_\varphi, e_\varphi\rangle}} = \frac{r^2\dot{\varphi}}{\sqrt{E}r} = \frac{p_\varphi}{\sqrt{E}r}.$$

故而 $r\cos\ \psi = \dfrac{p_\varphi}{\sqrt{E}} =$ 常数.

最后得到

定理 31.4.(克莱罗) 量 $r\cos\ \psi$ 沿三维空间 $\mathbb{R}^3$ 中曲面上的测地线保持不变.

因为 $p_\varphi = r^2\dot{\varphi} =$ 常数和 $2E = g_{zz}\dot{z}^2 + r^2\dot{\varphi}^2 =$ 常数, 我们得到 $r\dot{\varphi} = \dfrac{p_\varphi}{r}, 2E = g_{zz}(z)\dot{z}^2 + \dfrac{p_\varphi^2}{r^2}$, 其中 $r = r(z)$. 这允许对于旋转曲面上的测地线方程

$$d\varphi = \frac{p_\varphi}{r^2(z)}dt, \quad dt = \frac{dz}{\sqrt{\frac{2E}{g_{zz}} - \frac{p_\varphi^2}{g_{zz}r^2(z)}}} \tag{15}$$

进行完全积分.

习题

31.1. 设在拉格朗日函数 L 中包含了高阶导数:

$$L = L(t, x, \dot{x}, \ddot{x}, \cdots, x^{(k)}); \quad S(\gamma) = \int_\gamma L(t, x, \cdots, x^{(k)})dt.$$

证明对泛函 $S[\gamma]$ 的极值曲线成立下面的欧拉 – 拉格朗日方程:

$$\frac{\delta S}{\delta x} = \frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} + \frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{x}} - \cdots + (-1)^k\frac{d^k}{dt^k}\frac{\partial L}{\partial x^{(k)}} = 0.$$

§32. 守恒定律

1. 保持某个变分问题不变的变换群

利用单参数变换群 (见 §23), 可以对上一节所得到的动量守恒定律给出更加方便的不变形式.

设在 $\mathbb{R}^n$ 中给出了 (局部的) 单参数变换群 $S_\tau, -\infty < \tau < \infty$, 它具有下列性质:

1) 对任意点 P 存在数 $\tau_0 > 0$ 和点 P 在 $\mathbb{R}^n$ 中的邻域 U 使得当 $|\tau| < \tau_0$ 时 S_τ 有定义 (并且光滑):

$$S_\tau : U \to \mathbb{R}^n. \tag{1}$$

2) $S_0 = 1$(恒同变换),

$$S_{\tau_1+\tau_2} = S_{\tau_1} \circ S_{\tau_2}; \quad S_{-\tau} = S_\tau^{-1} \tag{2}$$

(在所有这些变换有定义处成立)

我们回忆, 每个局部单参数群有着相伴的向量场, 它切于轨线 $S_\tau(x)$:

$$(X^i) = X(x) = \frac{d}{d\tau} S_\tau(x)\bigg|_{\tau=0}. \tag{3}$$

反之, 向量场 (X^i) 也定义了群 S_τ, 它根据常微分方程解的存在和唯一性定理而得到. 这时对任意使 (X^i) 不为零的点 $x \in \mathbb{R}^n$ 可以找到这个点的邻域 U 及其上的坐标 $y^1, \cdots, y^n$, 使得变换 S_τ 当 τ 小时为

$$S_\tau(y^1, \cdots, y^n) = (y^1 + \tau, y^2, \cdots, y^n). \tag{4}$$

对这种形式的解的存在性定理在适当的地方还会被进一步应用.

定义 32.1. 称单参数变换群 S_τ 为拉格朗日函数 $L(x, \xi, t)$ 守恒是说

$$\frac{d}{d\tau} L(S_\tau(x), S_{\tau *}\xi, t) = 0, \tag{5}$$

其中 $S_{\tau *}$ 为切空间的映射 (见 §22.2).

令 $x(\tau) = S_\tau(x)$(使 $x = x(0)$), $\xi(\tau) = S_{\tau *}\xi(\xi = \xi(0))$, 方程 (4) 的左端便有形式

$$\frac{dL}{d\tau} = \dot{x}^i(\tau)\frac{\partial L}{\partial x^i(\tau)} + \frac{\partial L}{\partial \xi^i(\tau)}\frac{d\xi^i(\tau)}{d\tau}.$$

然而根据 §22.2 有 $\xi^i(\tau) = \xi^j \dfrac{\partial x^i(\tau)}{\partial x^j}$ (将 $x^i(\tau)$ 看作是 $x^j(0)$ 的函数, $j = 1, \cdots, n$), 故

$$\frac{dL}{d\tau} = \dot{x}^i(\tau)\frac{\partial L}{\partial x^i(\tau)} + \frac{\partial L}{\partial \xi^i(\tau)}\xi^j \frac{\partial}{\partial x^j}\left(\frac{d\xi^i(\tau)}{d\tau}\right).$$

计算当 $\tau = 0$ 时的导数, 我们便得到如下关系:

$$\frac{dL}{d\tau}\bigg|_{\tau=0} = X^i \frac{\partial L}{\partial x^i} + \frac{\partial X^i}{\partial x^j}\xi^j \frac{\partial L}{\partial \xi^i} = 0 \tag{6}$$

(函数 $L(x, \xi, t)$ 沿向量场 (X^i) 的李导数等于零). 这就是关于变换群 S_τ 为拉格朗日函数 L 守恒的条件.

有下面的定理.

定理 32.1. 如果单参数变换群 S_τ 为拉格朗日函数 L 守恒, 则成立动量分量沿向量场 $X=(X^i)$ 的守恒定律:

$$\frac{d}{dt}\left(X^i\frac{\partial L}{\partial \dot{x}^i}\right)=\frac{d}{dt}(X^ip_i)=0, \tag{7}$$

其中 $X(x)=\dfrac{d}{d\tau}S_\tau(x)|_{\tau=0}$.

证明 考虑点 $x\in\mathbb{R}^n$ 使 $X(x)\neq 0$, 并且有它的邻域 U, 在其中存在坐标 $(y^1,\cdots,y^n)$, 使在其下 S_τ 可写成上面所描述过的样子: $S_\tau(y^1,\cdots,y^n)=(y^1+\tau,y^2,\cdots,y^n)$, 在这样的坐标下拉格朗日函数守恒的条件为

$$\frac{\partial L(y,\dot{y})}{\partial y^1}\equiv 0.$$

因此在极值曲线 $y=y(t)$ 上有 $\dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{y}^1}\right)=0$. 但是沿轴 y^1 的单位向量是个向量场 X. 所以在所指出的坐标下 $\dfrac{\partial L}{\partial \dot{y}^1}=X^i\dfrac{\partial L}{\partial \dot{x}^i}$. 定理证完. □

2. 几个例子. 守恒定律的应用

例 32.1. 具非零质量的相对论中的 (自由) 粒子在具坐标 $(x^0,x^1,x^2,x^3), x^0=ct$ 的闵可夫斯基空间中定义了在类时曲线上的下面两个泛函 (作用) 中的一个:

$$S_1=\frac{mc}{2}\int\langle\dot{x},\dot{x}\rangle d\tau,\quad \langle\dot{x},\dot{x}\rangle=(\dot{x}^0)^2-\sum_{\alpha=1}^{3}(\dot{x}^\alpha)^2,\quad \dot{x}=\frac{dx}{d\tau}; \tag{8}$$

$$S_2=-mcl=-mc\int\sqrt{\langle\dot{x},\dot{x}\rangle}d\tau=-mc\int dl. \tag{9}$$

在空间 $\mathbb{R}^4_{1,3}$ 中的有质量的粒子的世界线是泛函 S_1 或 S_2 的极值曲线 (相对论的极小作用原理). 容易验证 (如同在 §31 中那样), 这两个泛函的极值曲线是一样的. 如同我们已看到的, 比照于经典力学时, 使用 S_2 是更为方便的.

在 (9) 的情形中,τ 可以取作任意的参数, 通常选取 $\tau=t=x^0/c$.

在此参数化下, 我们有

$$S_2=-mcl=-mc^2\int\sqrt{1-\frac{w^2}{c^2}}dt, \tag{10}$$

$$w^\alpha=\frac{dx^\alpha}{dt},\quad \alpha=1,2,3$$

($w=|w^\alpha|$ 为三维速度,l/c 为固有时间). 按通常的规则, 我们记 $S_2=\int Ldt$, 其中 $L=-mc^2\sqrt{1-\dfrac{w^2}{c^2}}$; 我们注意到这个拉格朗日函数 L 是个三维的量. 对这个拉格朗日函数的能量和动量有形式

$$E=pw-L=\dot{x}^\alpha\frac{\partial L}{\partial \dot{x}^\alpha}-L=\frac{mc^2}{\sqrt{1-\dfrac{w^2}{c^2}}}, \tag{11}$$

$$p_\alpha = \frac{mw^\alpha}{\sqrt{1-\frac{w^2}{c^2}}}. \tag{12}$$

当 $\frac{w}{c} \to 0$ 我们有

$$E \sim mc^2\left(1+\frac{w^2}{2c^2}+\cdots\right), \quad p_\alpha \sim mw^\alpha(1+\cdots),$$

即在经典表达式的一阶近似上我们得到动量为

$$p_\alpha \approx mw_\alpha \tag{13}$$

(参看 §31), 而准确到常数 mc^2 下, 能量就是经典表达式

$$E \approx mc^2 + \frac{mw^2}{2}. \tag{14}$$

此外, 还成立恒等式

$$E^2 - c^2p^2 = m^2c^4, \quad E = c\sqrt{p^2+m^2c^2}. \tag{15}$$

如果 $E > 0$, 则点 (E, cp) 取在闵可夫斯基空间 $\mathbb{R}_1^4$ 中的三维罗巴切夫斯基空间 (质量曲面) 上, 而 $\mathbb{R}_1^4$ 的坐标为 E, cp.

在此三维形式体系的框架中能量和动量是全然不同的, 而时间是一个单独的坐标.

我们转向泛函表示的作用 $S_1 = \frac{mc}{2}\int\left\langle\frac{dx}{d\tau}, \frac{dx}{d\tau}\right\rangle d\tau$. 由于上一节的结果, 在这里极值曲线的参数是自然参数 ("能量" 体系的守恒定律). 我们定义 4 动量 $\widetilde{p}_i = (\widetilde{p}_0, \widetilde{p}_\alpha), \alpha = 1, 2, 3$, 为

$$\widetilde{p}_i = \frac{\partial L}{\partial x'^i}, \quad i = 0, 1, 2, 3, \quad 其中\ x'^i = \frac{dx^i}{d\tau}.$$

我们得到

$$\widetilde{p}_0 = \frac{\partial L}{\partial x'^0} = mcx'^0, \quad \widetilde{p}_\alpha = \frac{\partial L}{\partial x'^\alpha} = -mcx'^\alpha, \quad \alpha = 1, 2, 3.$$

如果在闵可夫斯基度量下提升余向量 $\widetilde{p}_i$ 的指标, 则得到向量

$$\widetilde{p}^i = mcx'^i, \quad i = 0, 1, 2, 3.$$

因为沿泛函的极值曲线, 参数是自然的, 故而

$$dl = \sqrt{1-\frac{w^2}{c^2}}dt, \quad x'^i = \frac{dx^i}{dl},$$

$$\widetilde{p}^0 = mcx'^0 = \frac{mc^2}{\sqrt{1-\frac{w^2}{c^2}}} = E, \tag{16}$$

$$\widetilde{p}^\alpha = mcx'^\alpha = mc\frac{dx^\alpha}{dt} \cdot \frac{1}{\sqrt{1-\frac{w^2}{c^2}}} = \frac{mcw^\alpha}{\sqrt{1-\frac{w^2}{c^2}}} = cp^\alpha.$$

结论 4 动量由关系式

$$\widetilde{p}^{\,0} = E, \quad \widetilde{p}^{\,\alpha} = cp^\alpha \tag{17}$$

与三维能量和动量相关联. 也称 4 元动量为能量–动量向量.

结论 在洛伦兹变换下, 能量 动量向量 (E, cp) 如同 4 向量那样变化. 有质量的粒子的 4 动量位于有质量的曲面上. 曲面上具有罗巴切夫斯基几何结构:

$$E^2 - c^2p^2 = (\widetilde{p}^0)^2 - \sum_{\alpha=1}^{3}(\widetilde{p}^\alpha)^2 = m^2c^4, \tag{18}$$

特别, 对转移到以速度 v, 方向沿 x^1 轴作匀速运动的坐标系中, 我们有

$$(E, cp) \to (E', cp'),$$

其中

$$E' = \frac{E - p_1 v}{\sqrt{1-\frac{v^2}{c^2}}},$$

$$p_1' = \frac{-\frac{Ev}{c^2} + p_1}{\sqrt{1-\frac{v^2}{c^2}}}, \quad p_2' = p_2, \quad p_3' = p_3. \tag{19}$$

当 $\frac{v}{c} \to 0$, 我们得到 $E' = E - p_1 v +$ 常数,$p' = p +$ 常数.

例 32.2. 广义相对论的基本公设是爱因斯坦假设: 引力场不是别的, 而是四维时空空间中度量 g_{ij}(在每点的符号为 $(+---)$). 在没有其他任何力的情形下, 具质量 $m > 0$ 的测试粒子在引力场中按类时测地线运动, 此测地线由作用

$$S_1 = \frac{mc}{2}\int \langle \dot{x}, \dot{x}\rangle d\tau \tag{20}$$

的极值化定义.$m = 0$ 的粒子则按迷向测地线运动, 沿此曲线有 $\langle \dot{x}, \dot{x}\rangle = 0$.

由经典的引力势函数 $\varphi(x^0, x)$, $x^0 = ct$ 定义了弱场. 它由度量得出, 而此度量能形式地展开为 $1/c$ 的级数:

$$g_{ab} = g_{ab}^{(0)} + c^{-2}g_{ab}^{(2)} + O\left(\frac{1}{c^3}\right); \quad g_{00}^{(2)} = 2\varphi(x^0, x), \tag{21}$$

其中 $g_{ab}^{(0)}=\begin{pmatrix}1 & & & 0\\ & -1 & & \\ & & -1 & \\ 0 & & & -1\end{pmatrix}$ 为闵可夫斯基度量，其中含 $1/c$ 的这一项缺失.

定义 32.2. 在弱场中的类时测地线 $x^a=x^a(t), t=\dfrac{x^0}{c}$, 如果 $\left|\dfrac{dx^a}{dt}\right|\ll c$, 被称作为慢的.

对自然参数 (固有时间), 在弱场中的慢测地线为

$$d\tau=\frac{dl}{c}=\sqrt{\frac{1}{c^2}g_{ab}\frac{dx^a}{dt}\frac{dx^b}{dt}}dt\quad\left(t=\frac{x^0}{c}\right)$$

或者

$$d\tau=\left[1+O\left(\frac{1}{c^2}\right)\right]dt. \tag{22}$$

因此在测地线方程

$$\frac{d^2x^a}{d\tau^2}+\Gamma^a_{bc}\frac{dx^b}{d\tau}\frac{dx^c}{d\tau}=0$$

中, 其参数为自然参数, 在相差 $O\left(\dfrac{1}{c^2}\right)$ 下可以用 dt 替换 $d\tau$.

命题 32.1. 弱场中的慢测地线方程具形式

$$\frac{d^2x^\alpha}{dt^2}=-\frac{\partial\varphi}{\partial x^\alpha}+O\left(\frac{1}{c}\right), \tag{23}$$

即与在经典引力场的位势函数 φ 的牛顿方程相同.

证明 对于由公式

$$\Gamma^a_{bc}=\frac{1}{2}g^{ad}\left(\frac{\partial g_{bd}}{\partial x^c}+\frac{\partial g_{cd}}{\partial x^b}-\frac{\partial g_{bc}}{\partial x^d}\right)$$

进行计算的克里斯托费尔符号, 我们有:1) 由于 (21), 导数 $\dfrac{\partial g_{ab}}{\partial x^c}=\dfrac{1}{c}\dfrac{\partial g_{ab}}{dt}$ 的阶为 $O\left(\dfrac{1}{c^3}\right)$ (度量 $g_{ab}^{(0)}$ 为常值, 量 t 为有限);2) 导数 $\dfrac{\partial g_{ab}}{\partial x^\alpha}$ $(\alpha=1,2,3)$ 的阶为 $O\left(\dfrac{1}{c^2}\right)$(假定 x^α 为有限). 故而由在测地方程中具有克氏符号的那些项, 其阶最高的项是 $\Gamma^\alpha_{00}\dot{x}^0\dot{x}^0=\Gamma^\alpha_{00}c^2+O\left(\dfrac{1}{c}\right)$. 对于 Γ^α_{00} 我们有

$$\Gamma^\alpha_{00}=\frac{1}{2}g^{\alpha\alpha}\left(-\frac{\partial g_{00}}{\partial x^\alpha}\right)+O\left(\frac{1}{c^3}\right)=\frac{1}{c^2}\frac{\partial\varphi}{\partial x^\alpha}+O\left(\frac{1}{c^3}\right), \alpha=1,2,3.$$

由此得到所要的命题. □

例 32.3. n 个质点相互作用的经典系统由拉格朗日函数 (在空间 $\mathbb{R}^{3n}$ 中)

$$L = \sum_{i=1}^{n} \frac{m_i \dot{x}_i^2}{2} - U(x_1, \cdots, x_n), \quad x_i = (x_i^1, x_i^2, x_i^3) \tag{24}$$

给出, 其中 $U = \frac{1}{2}\sum_{i \neq j} V(x_i, x_j)$.

我们假定此系统为平移不变. 换句话说, 在变换

$$x_i \to x_i + \xi, \quad \xi = (\xi^1, \xi^2, \xi^3) \tag{25}$$

下拉格朗日函数不变: $L \to L$. 为此只要 V 为变量差的函数即可:

$$V(x_i, x_j) = V(x_i - x_j). \tag{26}$$

于是由动量守恒定律得到总动量守恒:

$$\frac{dP_{\text{II}}}{dt} = 0, \quad \text{其中} P_{\text{II}} = \sum_{i=1}^{n} m_i \dot{x}_i. \tag{27}$$

证明 我们有作用在 $\mathbb{R}^{3n}$ 上的三个群 S_τ^α ($\alpha = 1, 2, 3$), 其规则是

$$S_\tau^\alpha : x_i^\alpha \to x_i^\alpha + \tau, \quad x_i^\beta \to x_i^\beta, \quad \beta \neq \alpha, \quad \alpha = 1, 2, 3, \quad \beta = 1, 2, 3. \tag{28}$$

向量场 $X^{(\alpha)} = \left.\frac{d}{d\tau} S_\tau^\alpha(x)\right|_{\tau=0}$ 为

$$\begin{aligned} X^{(1)} &= (1, 0, 0, 1, 0, 0, \cdots, 1, 0, 0), \\ X^{(2)} &= (0, 1, 0, 0, 1, 0, \cdots, 0, 1, 0), \\ X^{(3)} &= (0, 0, 1, 0, 0, 1, \cdots, 0, 0, 1). \end{aligned} \tag{29}$$

根据一般情形的定理 32.1, 我们得到三个守恒定律:

$$P_{\text{II}, \alpha} = \sum_{i=1}^{n} m_i \dot{x}_i^\alpha, \quad \alpha = 1, 2, 3.$$

□

我们考虑 $n = 2, L = \frac{m_1(\dot{x}_1)^2}{2} + \frac{m_2(\dot{x}_2)^2}{2} - V(x_1 - x_2)$ 的情形. 我们转向匀速运动的坐标系, 在其中 $P_{\text{II}} = 0$. 于是有

$$m_1 \dot{x}_1 + m_2 \dot{x}_2 = 0. \tag{30}$$

我们选择质心作为原点, 于是 $m_1 x_1 + m_2 x_2 = 0$. 因此,

$$x_2 = -\frac{m_1 x_1}{m_2}, \quad V(x_1 - x_2) = V(x_1 + \frac{m_1}{m_2} x_1).$$

又牛顿方程为

$$m_1\ddot{x}_1^\alpha = -\frac{\partial V(x_1-x_2)}{\partial x_1^\alpha}, \quad \alpha = 1,2,3. \tag{31}$$

设 $m^* = m_1\left(1+\frac{m_1}{m_2}\right), U(x_1) = V\left(x_1+\frac{m_1}{m_2}x_1\right)$. 由 (31) 得到 (在与质心相关的坐标系中)

$$m^*\ddot{x}_1^\alpha = -\frac{\partial U(x_1)}{\partial x_1^\alpha}. \tag{32}$$

因此成立

定理 32.2. 两个质点在质心的坐标系下具有平移不变位势时的运动问题等价于一个具诱导质量 $m^* = m_1\left(1+\frac{m_2}{m_1}\right)$ 的质点在具位势 $U(x_1) = V(x_1-x_2)$ 的场中的运动, 其中 $m_1x_1+m_2x_2=0$.

因而平稳群引向到总动量的守恒, 并将两个质点的问题化为单个质点的问题.

例 32.4. 转而运用旋转群 $SO(3)$.

定义 32.3. 如果拉格朗日函数 $L(x,\dot{x})$ 对 $\mathbb{R}^3$ 中所有旋转不变, 则称其为球面对称.

在群 $SO(3)$ 中我们有三个不同的单参数子群:

1) 以角 φ 绕 x 轴的旋转 $(S_\varphi^{(x)})$;

2) 以角 φ 绕 y 轴的旋转 $(S_\varphi^{(y)})$;

3) 以角 φ 绕 z 轴的旋转 $(S_\varphi^{(z)})$.

它们在 $\mathbb{R}^3$ 中对应的向量场是在 §24.3 中找出的线性向量场 L_x, L_y, L_z:

$$L_x = (0,-z,y), \quad L_y = (z,0,-x), \quad L_z = (-y,x,0). \tag{33}$$

对于球面对称的拉格朗日函数, 根据一般性的定理 32.1 我们有下列量的守恒定律:

$$M_x = L_x^\alpha p_\alpha, \quad M_y = L_y^\alpha p_\alpha, \quad M_z = L_z^\alpha p_\alpha$$

(这是绕 x,y,z 轴旋转的动量的角分量). 它们的显式表达式为

$$\begin{gathered} M_x = yp_z - zp_y, \quad M_y = zp_x - xp_z, \quad M_z = xp_y - yp_x; \\ (\dot{M}_x) = (\dot{M}_y) = (\dot{M}_z) = 0 \end{gathered} \tag{34}$$

因此向量

$$M = (M_x, M_y, M_z) = [\bar{x}, p] \tag{35}$$

守恒, 其中 $\bar{x} = (x,y,z), p = (p_x,p_y,p_z)$. 向量 $[\bar{x},p]$ 称为动量矩.

设有一个两个质点的系统, 它对所有欧氏空间 $\mathbb{R}^3$ 的运动群不变. 这时 $L = \frac{m_1\dot{x}_1^2}{2} + \frac{m_2\dot{x}_2^2}{2} - U(x_1-x_2)$, 其中 $U(x_1-x_2) = U(|x_1-x_2|) = U(r)$. 事实上, 由

对所有旋转和反称的不变性得出 $U(x_1-x_2)$ 为 $r=|x_1-x_2|$ 的函数. 利用平稳群 (例 32.3 的结果), 转化为具质量 $m^*=m$ 的单个质点的问题, 此质点在具对稳群 $SO(3)$ 的场 $U(r)$ 中. 这时我们便有了动量矩 $M=[x,p]$ 的守恒定律. 因为 U 不依赖于时间, 则也使能量 $E=pv-L=\frac{mv^2}{2}+U(r), v=\dot{x}$ 守恒.

引理 32.1. 上述质点的运动只在由向量 x 和 p 张成的平面上进行.

证明 因为动量矩守恒以及 $\dot{x}=p/m$, 则向量 $[x,\dot{x}]=M/m$ 的方向不变. 此方向垂直于平面 (x,p). 证完. □

我们选取 z 轴为 M 的方向, 并使用柱坐标 (x,r,φ). 我们得到了

$$L=\frac{m\dot{x}^2}{2}-U(r)=m\frac{\dot{r}^2+r^2\dot{\varphi}^2}{2}-U(r),$$

$$p_\varphi=\frac{\partial L}{\partial\dot{\varphi}}=mr^2\dot{\varphi}=M=\text{常数},\quad \dot{\varphi}=\frac{M}{mr^2},\tag{36}$$

$$E=\frac{m}{2}\left(\dot{r}^2+\frac{M^2}{m^2r^2}\right)+U(r),$$

或者

$$E=\frac{m\dot{r}^2}{2}+U_3(r),\tag{37}$$

其中 $U_3(r)=U(r)+\frac{M^2}{2mr^2}$.

最后的结论是: 质点的运动化为 (对 r 的) 一维问题, 其势能为 $U_3(r)$; 其解的公式为

$$\dot{r}^2=\frac{2}{m}(E-U_3),\quad t-t_0=\int\frac{dr}{\sqrt{\frac{2}{m}(E-U_3)}},$$

$$\varphi-\varphi_0=\int\frac{M}{mr^2}dt.\tag{38}$$

消去 t, 可以得到轨线 $\varphi=\varphi(r)$ 或 $r=r(\varphi)$ 的方程.

在两个重要的已知情形 $U=\alpha/r$ 和 $U=\alpha r^2$ 中, 闭轨线充满了空间 $(x,\dot{x})$ 中下面的整个区域:

对 $U=\frac{\alpha}{r}, \alpha<0$, 区域为 $E<0$;

对 $U=\alpha r^2, \alpha>0$, 区域为整个 $E\geqslant 0$.

(在对 $U=\alpha/r$ 的区域 $E<0$ 是开普勒椭圆区域.) 闭轨要求满足非平凡的等式

$$r(\varphi+2\pi n)\equiv r(\varphi),\tag{39}$$

其中 n 为某个整数.

对于另外的球面对称的解析位势 $U(r)\neq\frac{\alpha}{r},\alpha r^2$, 等式 (39) 不能在相空间的整

个区域中成立 (参看书 [44]). 一般说来, 闭轨充满的集合其测度为零. 如果不等式

$$E - U_3(r) \leqslant 0$$

对有限线段 $0 \leqslant r_{\min} \leqslant r \leqslant r_{\max} < \infty$ 成立, 并且其中包含了初始条件, 则这个运动当 $-\infty < t < \infty$ 时发生在空间 $(x, \dot{x})$ 的有限区域中 (这里的 $U_3 = U(r) + \dfrac{M^2}{2mr^2}$, M^2 为总动量矩的平方).

例 32.5. 在闵可夫斯基空间 $\mathbb{R}^4$ 中给出了拉格朗日函数 $L(x, \dot{x}), x = (x^0, x^1, x^2, x^3)$, 它在洛伦兹群 $O(3,1)$ 下不变.$\mathbb{R}^4$ 中的群 $O(3,1)$ 对应于线性向量场, 而此向量场由单参数子群

$$X^i(x) - x_k A^{ki} = g^{(0)}_{kl} x^l A^{ki} \tag{40}$$

定义, 其中的 $g^{(0)}_{kl} = \begin{pmatrix} 1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & -1 \end{pmatrix}$ 为闵可夫斯基度量,A^{ki} 为任意的常值反称矩阵 (参看 §24). 于是我们有守恒定律

$$\widetilde{p}_i X^i = \widetilde{p}_i x_k A^{ki} = 常数, \tag{41}$$

其中的 $\widetilde{p}_i = \dfrac{\partial L}{\partial \dot{x}^i}$ 为 4 元动量. 由于矩阵 A^{ki} 的反对称性, 表达式 (41) 可以重写为

$$\widetilde{p}_i x_k A^{ki} = \frac{1}{2}(\widetilde{p}_i x_k - \widetilde{p}_k x_i) A^{ki} = 常数.$$

因为矩阵 A^{ki} 是任意的, 故我们有了保持上面的量不变的整个张量

$$M_{ik} = x_i \widetilde{p}_k - x_k \widetilde{p}_i = 常数 \tag{42}$$

称这个张量为*4 元动量矩张量*.

为方便起见, 我们考虑相应的上指标张量 (M^{ik}). 它的空间坐标分量 $M^{\alpha\beta}, \alpha, \beta = 1, 2, 3$ 有形式

$$M^{\alpha\beta} = x^\alpha \widetilde{p}^\beta - x^\beta \widetilde{p}^\alpha, \tag{43}$$

即与三维动量矩向量的分量

$$cM = c[x, p], \quad M^{23} = cM_x, \quad M^{31} = cM_y, \quad M^{12} = cM_z \tag{44}$$

重合. 分量 M^{01}, M^{02}, M^{03} 则构成了三维向量

$$(M^{01}, M^{02}, M^{03}) = c^2 tp - Ex. \tag{45}$$

公式 (44) 和 (45) 立即由 4 元动量

$$(\widetilde{p}^i) = (E, cp)$$

得到, 这里的 p 为通常的三维动量.

我们现在设给出了由 n 个相对论的粒子 $x_1, \cdots, x_n$ 构成的系统, 其中 $x_i = \{x_i^0, x_i^1, x_i^2, x_i^3\}$. 进一步, 设拉格朗日函数 $L(x_1, \cdots, x_n, \dot{x}_1, \cdots, \dot{x}_n)$ 在整个庞加莱群下不变; 我们知道, 庞加莱群是闵可夫斯基空间的运动群. 于是我们有下面的守恒定律:

$$\left(\sum_{i=1}^{n} E_i, \sum_{i=1}^{n} cp_i\right) = \text{常数 (总 4 元动量守恒)},$$

$$\sum_{i=1}^{n} M_i^{kl} = \text{常数 } (M_i^{kl}\text{为第 } i \text{ 个质量的动量矩张量}).$$

由公式 (45) 特别得到

$$\sum c^2 t p_i - \sum E_i x_i = \text{常数}. \tag{46}$$

因为总能量 $\sum E_i$ 也守恒, 则有

$$\frac{\sum E_i x_i}{\sum E_i} = t \cdot \frac{\sum c^2 p_i}{\sum E_i} + \text{常数}.$$

从而点

$$x = \frac{\sum E_i x_i}{\sum E_i} \tag{47}$$

以常速 v 运动, 其中

$$v = \frac{\sum c^2 p_i}{\sum E_i}. \tag{48}$$

点 x 为相对论的类比质心. 如果粒子的速度相比于 c 是小的, 则可近似地令 $E_i \approx m_i c^2$, 从而公式 (4 7) 转化为经典的公式

$$x = \frac{\sum m_i x_i}{\sum m_i}.$$

我们注意到, 相对论的质心对于参照系的选取是不变的.

§33. 哈密顿体系

1. 勒让德变换

我们记得曾经令

$$E = \dot{x}^\alpha \frac{\partial L}{\partial \dot{x}^\alpha} - L, \quad p_\alpha = \frac{\partial L}{\partial \dot{x}^\alpha} \tag{1}$$

从而引进了对于拉格朗日 (函数)$L(x, \dot{x})$ 的能量和动量.

定义 33.1. 如果

$$\det\left(\frac{\partial^2 L}{\partial \dot{x}^\alpha \partial \dot{x}^\beta}\right) \neq 0, \tag{2}$$

则称 L 为非异的, 如果方程 $p_\alpha = \dfrac{\partial L(x, \dot{x})}{\partial \dot{x}^\alpha}$ 可以光滑且一一地解出为 $\dot{x}^\alpha = v^\alpha(x, p)$ 的形式, 其中 $x, \dot{x}$ 任意, 则称其为强非异的.

定义 33.2. 称强非异拉格朗日的由 x, p 表示的能量 E 为哈密顿 $H(x, p)$.

称其坐标 (x, p) 的空间为相空间. 对于强非异拉格朗日, 我们可以从坐标 $(x, \dot{x})$ 转移到坐标 (x, p). 这是由 $L(x, v)$ 到 $H(x, p)$ 的勒让德变换. 这时欧拉 – 拉格朗日方程转换成了哈密顿方程.

定理 33.1. 设 $L(x, \dot{x})$ 为强非异拉格朗日,$H(x, p)$ 为哈密顿, $p = \dfrac{\partial L(x, \dot{x})}{\partial \dot{x}}, \dot{x} = v(x, p)$. 欧拉 – 拉格朗日方程

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = 0 \tag{3}$$

等价于下面的哈密顿方程, 在其中 x 和 p 被假设为相空间 (x, p) 的独立变量:

$$\dot{p} = -\frac{\partial H}{\partial x}, \quad \dot{x} = \frac{\partial H}{\partial p}. \tag{4}$$

证明 设 $L = L(x, v), v = \dot{x}$,

$$H = pv - L, \quad v = v(x, p), \quad p = \left.\frac{\partial L}{\partial v}\right|_{x=\text{常数}}.$$

我们要证明下面的等式

a) $\dot{x} = v = \left.\dfrac{\partial H}{\partial p}\right|_{x=\text{常值}}$,

b) $\dot{p} = -\left.\dfrac{\partial H}{\partial x}\right|_{p=\text{常值}} = \left.\dfrac{\partial L}{\partial x}\right|_{v=\text{常值}}$,

其中 $H = H(x, p), L = L(x, v)$.

等式 a) 的证明: 因为 $p = \dfrac{\partial L}{\partial v}$, 故

$$\frac{\partial H}{\partial p}=\frac{\partial}{\partial p}(pv-L)=v+p\frac{\partial v}{\partial p}-\frac{\partial L}{\partial v}\frac{\partial v}{\partial p}=v.$$

等式 b) 的证明:

$$-\frac{\partial H}{\partial x}=-\frac{\partial}{\partial x}(pv-L)=-p\frac{\partial v}{\partial x}+\frac{\partial L}{\partial x}+\frac{\partial L}{\partial v}\frac{\partial v}{\partial x}=\frac{\partial L}{\partial x}=\dot{p}.$$

证完. □

我们注意到 $L=pv-H$, 其中 $v=\dfrac{\partial H}{\partial p}$, 或者 $L=p\dot{x}-H$. 对 $L(x,\dot{x})$, 作用 $S=\int Ldt$ 被写为

$$S=\int Ldt=\int[p\dot{x}-H(x,p)]dt. \tag{5}$$

如果曲线 γ 为 $x(t)$ 和 $p(t)=\dfrac{\partial L}{\partial \dot{x}}(x,\dot{x})$, 则沿此曲线满足 "可积性条件"

$$v(t)=\frac{dx}{dt},\quad 其中\ v=\frac{\partial H}{\partial p}. \tag{6}$$

我们考虑在相空间 (x,p) 中拉格朗日 $L=p\dot{x}-H(x,p)$, 它在曲线 $x(t),p(t)$ 上不满足可积性条件 $\dfrac{dx}{dt}=\dfrac{\partial H}{\partial p}$.

引理 33.1. 对于在 $2n$ 维空间 (x,p) 中的泛函 $\int[p\dot{x}-H(x,p)]dt$, 其欧拉 – 拉格朗日方程与哈密顿方程相重合, 并且自动地蕴涵了可积性条件 $v=\dfrac{\partial H}{\partial p}=\dfrac{dx}{dt}$.

证明 设 $\widetilde{L}(y,\dot{y})=p\dot{x}-H(x,p)$, 其中 y 为坐标 (x,p). 如 $y^\alpha=x^i$, 则欧拉 – 拉格朗日方程 $\dfrac{d}{dt}\left(\dfrac{\partial\widetilde{L}}{\partial\dot{y}^\alpha}\right)=\dfrac{\partial\widetilde{L}}{\partial y^\alpha}$ 具有形式

$$\frac{d}{dt}\left(\frac{\partial\widetilde{L}}{\partial\dot{x}^i}\right)=\frac{\partial\widetilde{L}}{\partial x^i}\ 或\ \dot{p}=-\frac{\partial H}{\partial x},$$

而如果 $y^\alpha=p_i$, 则

$$0=\frac{d}{dt}\left(\frac{\partial\widetilde{L}}{\partial\dot{p}_i}\right)=\frac{\partial\widetilde{L}}{\partial p_i}=\dot{x}^i-\frac{\partial H}{\partial p_i}.$$

引理得证. □

注 在上面我们推导出了对于在固定端点的曲线类中成为极值曲线的欧拉 – 拉格朗日方程. 在 (x,p) 相空间中它对应于一种不同的曲线类, 它只固定了端点的 x 坐标 (图 34).

2. 活动坐标系

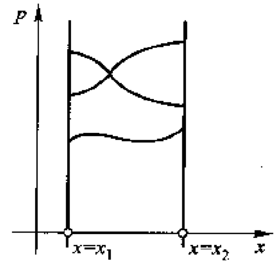


图 34

我们现在考虑的情形是在 $\mathbb{R}^n$ 中给出了随时间流而变化的拉格朗日 $L=L(x,\dot{x},t)$. 我们将推导出在变换 $x=x(x',t), t=t'$ 下不同量之间的变化规则.

首先, 按定义, 在此变换下 $S=\int Ldt$ 没有变化, 而 L 也没有变化 (仅差一个全导数). 因此我们有

$$L(x,\dot{x},t)\to L(x(x',t),\dot{x},t)+\frac{df(x,t)}{dt}. \qquad (7)$$

考虑两种情形:

1) 不含时间的变换:$x=x(x')$;

2) 含时间的变换.

在第一种情形,

$$\dot{x}^i=\frac{\partial x^i}{\partial x^{i'}}\dot{x}^{i'} \quad \text{或者} \quad v^i=\frac{\partial x^i}{\partial x^{i'}}v^{i'}. \qquad (8)$$

结论 速度是个向量.

另外,$\widetilde{L}(x',\dot{x}',t)=L(x,\dot{x},t)$. 所以

$$\widetilde{L}(x',\dot{x}',t)=L(x(x'),\frac{\partial x}{\partial x'}v',t),$$

$$p_{i'}=\frac{\partial\widetilde{L}}{\partial v^{i'}}=\frac{\partial L}{\partial v^i}\frac{\partial x^i}{\partial x^{i'}}=p_i\frac{\partial x^i}{\partial x^{i'}}, \qquad (9)$$

就是说, 动量是个余向量. 对于能量我们有 $E=p_iv^i-L=p_{i'}v^{i'}-\widetilde{L}=E'$. 因此能量是个标量 (无变化).

现在转而考虑含时间的变换:

$$x=x(x',t), \quad t=t'. \qquad (10)$$

此时

$$\dot{x}^i=\frac{\partial x^i}{\partial x^{i'}}\dot{x}^{i'}+\frac{\partial x^i}{\partial t}=\frac{\partial x^i}{\partial x^{i'}}\dot{x}^{i'}+a^i(x^i,t). \qquad (11)$$

我们假定, 在给定的时刻 $t=t_0$, 构造了 (不依赖时间的) 变换, 使当 $t=t_0$ 时 $x=x'$. 于是

$$v^i=v'^i+\frac{\partial x^i}{\partial t}=v'^i+a^i(x',t_0), \quad t_0=t. \qquad (12)$$

称这样的坐标系为瞬时的. 因而在活动坐标系下我们得到

$$L\to L'=L, \quad L(x,v,t)=L(x',v'+a,t);$$

$$v\to v'=v-a(x',t);$$

$$p\to p'=p, \text{因为有 } \frac{\partial L}{\partial v}=\frac{\partial L'}{\partial v'};$$

$$E\to E'=p'v'-L'=p_i(v^i-a^i)-L=E-p_ia^i(x',t).$$

因此, 动量没有变化, 而能量移动了一个量, 它等于 $-p_i a^i = \langle p, v' - v\rangle$. 因此,

$$E = H(x,p) \to E' = H(x',p') - p'_i a^i(x',t)$$
$$= H'(x',p',t), \tag{13}$$

其中当 $t = t_0$ 时 $x = x', p = p', a = \left.\dfrac{\partial x}{\partial t}\right|_{t=t_0}$. 我们来解释哈密顿是如何变化的.

首先我们有 $\dot{x} = \dfrac{\partial H}{\partial p}, \dot{p} = -\dfrac{\partial H}{\partial x}$. 现在则有 $\dot{x}' = \dfrac{\partial H'}{\partial p'}, \dot{p}' = -\dfrac{\partial H'}{\partial x'}$(当 $t = t_0$ 时), 因为 $x' = x, p' = p$, 我们得到

$$\dot{x}' = \frac{\partial H'}{\partial p'} = \frac{\partial H}{\partial p} - a = \dot{x} - a,$$

$$\dot{p}' = -\frac{\partial H'}{\partial x'} = -\frac{\partial H}{\partial x} + p'_i \frac{\partial a^i}{\partial x'} = \dot{p} + p_i \frac{\partial a^i}{\partial x'}. \tag{14}$$

因此, 哈密顿的方程的变化十分简单. 我们已证明了下面的命题.

定理 33.2. 当在活动坐标系时, 瞬时坐标和动量没有变化, 而能量 (哈密顿) 改变了量 $-\langle p, a\rangle$

$$H \to H - \langle p, a\rangle, \quad a = a(x',t) = \left.\frac{\partial x}{\partial t}\right|_{t=t_0}.$$

考察二个重要的例子.

例 33.1. 坐标系的平移运动, 这时 $a = a(t)$ 不依赖于 $x', L = \dfrac{m\dot{x}^2}{2} - U(x)$. 此时

$$p = mv, \quad p' = p = mv = m(v' + a) = mv' + ma, \quad H' = H - \langle p', a\rangle.$$

对牛顿方程

$$(mv)^{\cdot} = f = m\dot{v}' + m\dot{a}$$

或者

$$m\dot{v}' = f - m\dot{a} = f'. \tag{15}$$

因此, 力 f 得到了一个惯性补充 $-m\dot{a}$.

例 33.2. $\mathbb{R}^3$ 中坐标系的旋转, 这里 $a = [\Omega, x'], \Omega$ 为角速度 (常向量);

$$H' = H - \langle p', a\rangle = H - \langle p', [\Omega, x']\rangle,$$

$$p' = p, x' = x, \text{这时 } t = t_0.$$

方程便为

$$\dot{p}' = -\frac{\partial H'}{\partial x'} = -\frac{\partial H}{\partial x} + \frac{\partial}{\partial x}\langle p', [\Omega, x]\rangle = -\frac{\partial H}{\partial x} + [p', \Omega],$$

$$\dot{x}' = \frac{\partial H'}{\partial p'} = \frac{\partial H}{\partial p} - \frac{\partial}{\partial p'}(p', [\Omega, x]) = \frac{\partial H}{\partial p} - [\Omega, x'].$$

如果拉格朗日为 $L = \frac{m\dot{x}^2}{2} - U(x)$, 则

$$p = p' = m\dot{x} = m(\dot{x}' + [\Omega, x']) = m\dot{x}' + ma.$$

我们便有

$$\dot{p}' = m\ddot{x}' + m[\Omega, x']^{\cdot} = -\frac{\partial H}{\partial x} + [(m\dot{x}' + ma), \Omega].$$

最后得到: 如果 $\dot{\Omega} = 0$(常速旋转), 则

$$m\ddot{x}' = 2m[\dot{x}', \Omega] + f + m[[\Omega, x], \Omega]. \tag{16}$$

f 为原来在固定坐标下的力. 最后的一项当 Ω 为小,$|x|$ 不大时是一个小的力, 称力 $2m[\dot{x}', \Omega]$ 为科里奥利 (Coriolis) 力.我们得到

$$m\dot{v}' = 2m[v', \Omega] + f + O(\Omega^2). \tag{17}$$

注 如果 $L = \frac{m\dot{x}^2}{2} - U(x)$, 转入活动坐标系中则有

$$H \to H - p_i a^i, \quad p_i = mv^i = p_i' = m(v'^i + a^i), \tag{18}$$

其中

$$H = \frac{p^2}{2m} + U(x), \tag{19}$$

$$H' = \frac{p'^2}{2m} + U(x) - pa = \frac{(p' - ma)^2}{2m} + U(x) - \frac{ma^2}{2}.$$

因此, 到活动坐标系的转移等价于下面的两个运算:

a) 动量的移动 $p \to p' - ma$,

b) 位势的变换 $U \to U_3 = U(x) - \frac{ma^2}{2}$.

例 33.3. (含电磁场).设 $L(x, \dot{x})$ 为拉格朗日. 我们来定义一个新的拉格朗日函数为

$$\widetilde{L} = L + \frac{e}{c}A_i\dot{x}^i,$$

其中 A_i 为电磁场的势向量,e 为质点的电荷量, $S = \int L dt, \widetilde{S} = \int (L dt + \frac{e}{c}A_i dx^i)$. 在拉格朗日函数上加上一项 $\frac{e}{c}A_i\dot{x}^i$ 的运算被称做 "含电磁场".

如果 $H(x, p) = pv - L$ 和 $\widetilde{H}(x, \widetilde{p}) = \widetilde{p}v - \widetilde{L}$, 其中 $p = \frac{\partial L}{\partial v}, \widetilde{p} = \frac{\partial \widetilde{L}}{\partial v}$, 于是

$$\begin{aligned}\widetilde{p}_i &= p_i + \frac{e}{c}A_i, \quad p_i = \widetilde{p}_i - \frac{e}{c}A_i, \\ \widetilde{H}(x,\widetilde{p}) &= H(x,\widetilde{p} - \frac{e}{c}A).\end{aligned} \tag{20}$$

因此, 场的内含等价于动量在哈密顿中的位移, 因而类似于转移到活动坐标系 (无势的变换).

我们注意到, 欧拉 – 拉格朗日方程为 $\dot{\widetilde{p}}_i = \dfrac{\partial \widetilde{L}}{\partial x^i} = \dfrac{\partial L}{\partial x^i} + \dfrac{e}{c}\dfrac{\partial A_j}{\partial x^i}\dot{x}^j$. 然而, 作为由 (20) 所得, 有 $\dot{\widetilde{p}}_i - \dot{p}_i + \dfrac{e}{c}\dfrac{\partial A_i}{\partial x^j}\dot{x}^j$. 故而

$$\begin{aligned}\dot{p}_i &= f_i = \frac{\partial L}{\partial x^i} \ (\text{无电磁场}), \\ \dot{p}_i &= f_i + \frac{e}{c}F_{ij}\dot{x}^j \ (\text{含电磁场}),\end{aligned} \qquad p = \widetilde{p} - \frac{e}{c}A.$$

这里的 $F_{ij} = \dfrac{\partial A_j}{\partial x^i} - \dfrac{\partial A_i}{\partial x^j}$ 为电磁场张量. 由张量 F_{ij} 的反称性, 表达式 $F_{ij}\dot{x}^i\dot{x}^j$ 恒为零. 因此

$$\dot{p}_i\dot{x}^i = \dot{p}_i v^i = f_i v^i, \tag{21}$$

其中 f 为无电磁场时的力, $f = \dfrac{\partial L}{\partial x}$.

注 在三维体系中分量 $A_0 = c\varphi$ 为电磁势能,(A_1, A_2, A_3) 为势向量. 在这些记号下, 拉格朗日的添加为 $e\varphi + \dfrac{e}{c}A_\alpha \dot{x}^\alpha$, 其中 $x^0 = ct, \alpha = 1, 2, 3$.

3. 莫佩尔蒂和费马原理. 应用.

我们现在来分析莫佩尔蒂和费马的变分原理. 在此之前, 我们来阐述在三维哈密顿中的能量守恒定律: 在 $2n$ 维相空间沿运动的轨线 $x(t), p(t)$(即沿泛函 $S = \int (p\dot{x} - H)dt$ 的极值线或哈密顿方程的解), 哈密顿 $H(x(t), p(t))$ 为常值. 此时成立

定理 33.3.(莫佩尔蒂原理) 如果哈密顿 $H(x,p)$ 与时间无关, 并且能量水平固定, 则在具所给能量 E 的上述所有曲线 $x(t)$ 中的极值曲线也是 $S_0 = \int p dx$ 的极值曲线. 此时我们假定形式 pdx 由 x^α, dx^α 表示.

证明 显然, 在曲面 $H(x,p) = E$ 上, 原来的泛函 $S = \int (pdx - Hdt)$ 在所有原来的极值线 (哈密顿系统的解) 上达到了极值, 这是因为我们是在较原来的更小的曲线类上考虑极值化问题: 这些曲线现在位于曲面 $H(x,p) = E$ 上. 在此曲面上,$S = \int (pdx - Edt) = \int (pdx - d(Et))$. 故而在此曲面 (维数为 $2n-1$) 的坐标 $z^1, \cdots, z^{2n-1}$ 下, 拉格朗日 $p\dot{x} = \widetilde{L}(z, \dot{z})$ 和 $\widetilde{L}(z, \dot{z}) - \dfrac{d}{dt}(Et)$ 有同样的极值曲线方程.

因此, 在曲面 $H=E$ 上, 新的 (水平的)$S_0=\int pdx$ 的极值线即是原来的在 (x,p) 空间中泛函的所有极值线. 定理得证. □

现在来考虑重要例子.

例 33.4. $L=\frac{m\dot{x}^2}{2}-U(x); H(x,p)=\frac{p^2}{2m}+U(x)$. 沿极值线满足方程

$$\dot{x}=\frac{\partial H}{\partial p}=\frac{p}{m}.$$

我们将水平作用的泛函 $S_0=\int pdx$ 限制在还要比曲面 $H=E$ 上曲线集合还要小的集合上, 即要求在这些曲线上要满足关系 $\dot{x}=p/m$. 于是:

a) 由条件 $H(x,p)=E$ 有 $|p|=\sqrt{2m(E-U(x))}$;

b) 由条件 $\dot{x}=p/m$ 有 $\langle p,\dot{x}\rangle=p_\alpha\dot{x}^\alpha=|p|\cdot|\dot{x}|$.

由此证明了

定理 33.4. 对应于在固定能量 E 时的哈密顿 $H=p^2/2m+U(x)=E$ 的哈密顿系统的解曲线 $x(t)$ 是条测地线, 其对应的新度量为

$$g_{ij}=2m(E-U(x))\delta_{ij} \tag{22}$$

(这时的参数并不是自然的).

例 33.5. 设 $H=c(x)\cdot|p|$; 这个哈密顿描述了具变化光速 $c(x)$ 的速向球面上的光的轨线.

考虑在一个曲线集合上的水平作用 $S_0=\int pdx$, 这个集合满足

a) $H(x,p)=E$,

b) $\dot{x}=\frac{\partial H}{\partial p}=c(x)\frac{p}{|p|}$ (显然,$|\dot{x}|=c(x)$).

因为 $c(x)|p|=E$, 故 $|p|=E/c(x)$. 因为 $\dot{x}=c(x)\frac{p}{|p|}$, 故 $\langle p,dx\rangle=|p|\cdot|dx|$. 从而

$$\begin{aligned}\int\langle p,dx\rangle&=\int|p|\cdot|dx|=\int\frac{E}{c(x)}|dx|=E\int\frac{|dx|}{c(x)}\\&=E\int\sqrt{\frac{dx^2}{c^2(x)}}=E\int\sqrt{g_{ij}\dot{x}^i\dot{x}^j}\,dt,\end{aligned}$$

其中 $g_{ij}=\frac{1}{c^2(x)}\delta_{ij}$. 我们注意, 积分 $\int_\gamma\frac{|dx|}{c(x)}$ 等于光沿路径 γ 的运动时间.

于是已经证明了

定理 33.5. (费马原理) 光沿着在连接所给点 P 和 Q 的所有光滑曲线中使运动时间达到极值的曲线运动. 曲线是在新度量

$$g_{ij}=\frac{1}{c^2(x)}\delta_{ij}$$

下的测地线.

注 测地线通常由拉格朗日 $L=g_{ij}v^iv^j$ 或哈密顿 $H=g^{ij}p_ip_j$ 得到. 如果取哈密顿为 $H'(x,p)=\sqrt{H}=\sqrt{g^{ij}p_ip_j}$, 则它将给出同样的运动轨线; 这是由于在常值水平能量 $E=$ 常数上, 相应的哈密顿 H 和 $\sqrt{H}$ 的向量场将以常数因子 $\dfrac{1}{2\sqrt{H}}=\dfrac{1}{2\sqrt{E}}=$ 常数成比例. 事实上,$(\dot{x},\dot{p})=\left(\dfrac{\partial H}{\partial p},-\dfrac{\partial H}{\partial x}\right)=2\sqrt{H}\left(\dfrac{\partial\sqrt{H}}{\partial p},-\dfrac{\partial\sqrt{H}}{\partial x}\right)$.

在我们的情形中 $H'=c(x)|p|$ 或 $H'=\sqrt{c^2(x)|p|^2}$. 故而立即可以指出, 度量具有形式 $g^{ij}=c^2(x)\delta^{ij}$ 或 $g_{ij}=\dfrac{1}{c^2(x)}\delta_{ij}$.

在各向异性的介质中, 定义了速度的张量 g_{ij}(或 g^{ij}) 不再共形于欧几里得形式.

§34. 相空间的几何理论

1. 梯度系统

设在任意空间 $\mathbb{R}^m$ 中坐标为 $(y^1,\cdots,y^m)$, 而 "度量"g^{ij} 可以定义余向量的内积, 并且可以提升指标 (参看 §19).

函数 $f(y^1,\cdots,y^m)$ 的梯度 ∇f 为

$$(\nabla f)^i=g^{ij}\frac{\partial f}{\partial y^j}. \tag{1}$$

这时张量 g^{ij} 并没有假定为对称的. 对应于向量场 ∇f 方程组 (参看 §23) 为

$$\dot{y}^i=(\nabla f)^i. \tag{2}$$

称形如 (2) 的系统为梯度. 有简单的引理:

引理 34.1. 设 $y(t)$ 为梯度方程组 (2) 的积分曲线. 于是对任意函数 $h(y)$, 其对 t 的导数为

$$\dot{h}-\frac{d}{dt}h(y(t))=\langle\nabla h,\nabla f\rangle=g^{ij}\frac{\partial h}{\partial y^i}\frac{\partial f}{\partial y^j}. \tag{3}$$

证明 由定义

$$\dot{h}=\frac{\partial h}{\partial y^i}\dot{y}^i=\frac{\partial h}{\partial y^i}(\nabla f)^i=\frac{\partial h}{\partial y^i}g^{ij}\frac{\partial f}{\partial y^j}.$$

引理证毕. □

使我们感兴趣的是非异的反称"度量"g^{ij}, 它由形式

$$\Omega=\frac{1}{2}g_{ij}dy^i\wedge dy^j \tag{4}$$

给出, 其中的 g_{ij} 为逆矩阵:$g_{ij}g^{jk}=\delta_i^k, g_{ij}=-g_{ji}, \det(g_{ij})=g\neq 0$. 显然, 空间的维数是偶数,$i,j=1,\cdots,2n$.

引理 34.2. 有公式

$$-\frac{1}{n!}\Omega^n = \frac{1}{n!}\underbrace{\Omega\wedge\cdots\wedge\Omega}_{n} = \sqrt{g}dy_1\wedge\cdots\wedge dy^{2n}, \tag{5}$$

因此, $\sqrt{g}$ 为 g_{ij} 的多项式; 称其为普法夫 (多项式).

证明 只要在每个点分别证明等式 (5) 就可以了. 在每个选取的点 P 上, 张量 g_{ij} 是在点 P 的切空间的反称形式. 因为与对称形式在适当基下取对角形式一样, 反称形式 $g_{ij}(P)$ 也可化为 $\begin{pmatrix} O_n & I_n \\ -I_n & O_n \end{pmatrix}$, 其中 O_n 和 I_n 分别是 $n\times n$ 阶的零矩阵和单位矩阵. 因此, 如果选取新坐标

$$(x^1,\cdots,x^n,p_1,\cdots,p_n) = (z^1,\cdots,z^{2n}), \tag{6}$$

使其与老的坐标线性相联系:$y = Az$, 并且使

$$(\widetilde{g}_{ij}(P)) = A^{\mathrm{T}}(g_{ij}(P))A = \begin{pmatrix} O_n & I_n \\ -I_n & O_n \end{pmatrix},$$

于是此时有

$$\Omega = \sum_i dx^i\wedge dp_i, \quad \Omega\wedge\cdots\wedge\Omega = -n!dz^1\wedge\cdots\wedge dz^{2n}. \tag{7}$$

因为 $\sqrt{\widetilde{g}} = 1$, 故而在所给坐标系中引理得证. 由于形式 (5) 的不变性 (这个关系式的右端在具正雅可比的变换下像张量一样地变化). 引理便在任意坐标系下得到了证明.(在 $\det A < 0$ 时其证明类似.) □

注 因此, 度量 g_{ij} 非异的条件 $g\neq 0$ 等价于条件 $\Omega^n\neq 0$.

定义 34.1. 在具反称度量 g_{ij} 的空间中如有坐标 (x,p) 使得$g_{ij}\equiv\begin{pmatrix} O_n & I_n \\ -I_n & O_n \end{pmatrix}$, 则称此空间为*相空间*.使度量为 (7) 的那个坐标 (x,p) 被称为*典则坐标*. 称在此度量下的梯度系统为*哈密顿系统*.

在典则坐标下的哈密顿系统(以函数 $H(x,p)$ 作为 f) 的形式为

$$\dot{y}^i = (\nabla H)^i, \quad i = 1,\cdots,2n,$$

或者

$$\dot{x}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial x^i}, \quad i = 1,\cdots,n. \tag{8}$$

读者不难在这些方程中认出哈密顿方程的样子. 哈密顿系统的重要性是由前一节中证明了的定理决定了的, 即关于欧拉 – 拉格朗日方程 (对于强非异拉格朗日) 和相空间中的哈密顿方程之间等价性的定理.

由引理 34.1 我们得到

引理 34.3. 任意函数 $f(x,p,t)$ 沿哈密顿系统 (8) 的轨线的导数为

$$\dot{f}=\frac{\partial f}{\partial t}+\langle\nabla f,\nabla H\rangle,\tag{9}$$

其中

$$\langle\nabla f,\nabla H\rangle=g^{ij}\frac{\partial f}{\partial y^i}\frac{\partial H}{\partial y^j}=\sum_{i=1}^{n}\left(\frac{\partial f}{\partial x^i}\frac{\partial H}{\partial p_i}-\frac{\partial H}{\partial x^i}\frac{\partial f}{\partial p_i}\right).\tag{10}$$

特别, 如果 $f=H=H(x,p,t)$, 由于 $\langle\nabla H,\nabla H\rangle=-\langle\nabla H,\nabla H\rangle=0$, 则 $\dot{E}=\dot{H}=\frac{\partial H}{\partial t}$.

现在考虑与 t 显式相关的系统 $H=H(x,p,t)$, 我们有哈密顿方程

$$\dot{x}=\frac{\partial H}{\partial p},\quad \dot{p}=-\frac{\partial H}{\partial x},\tag{11}$$

以及它的推论 $\dot{E}=\frac{\partial H}{\partial t}$ 和显然的等式 $\dot{t}=1$. 我们考虑一个扩张的相空间, 其坐标为 (x,p,t,E), 其中 $x^{n+1}=t, p_{n+1}=E$, 并且反称度量为

$$\widehat{g}_{ij}=\begin{pmatrix}0&1&&&&&\\-1&0&&&&&\\&&\ddots&&&&\\&&&0&1&&\\&&&-1&0&&\\&&&&&0&1\\&&&&&-1&0\end{pmatrix},\tag{12}$$

它对应的形式为

$$\widehat{\Omega}=\sum_{i=1}^{n}dx^i\wedge dp_i-dt\wedge dE.\tag{13}$$

考虑哈密顿 $\widetilde{H}(x,p,t,E)=H(x,p,t)-E$ 和梯度系统 (哈密顿), 它有形式

$$\dot{x}=\frac{\partial\widetilde{H}}{\partial p}=\frac{\partial H}{\partial p},\quad \dot{p}=-\frac{\partial\widetilde{H}}{\partial x}=-\frac{\partial H}{\partial x},\quad \dot{t}=-\frac{\partial\widetilde{H}}{\partial E}=1,\quad \dot{E}=\frac{\partial\widetilde{H}}{\partial t}=\frac{\partial H}{\partial t}.\tag{14}$$

由这些方程得出.

推论 在扩张的相空间 (x,p,t,E) 中, 其度量为 $\widehat{g}_{ij}$(或者 $\widehat{\Omega}$), 且哈密顿为 $\widetilde{H}=H(x,p,t)-E$; 则曲面 $\widetilde{H}=0$ 上的哈密顿方程与空间 (x,p) 中的原来的哈密顿方程再加上关系式 $\dot{t}=1, \dot{E}=\frac{\partial H}{\partial t}$ 相重合.

如果原来的哈密顿 H 并不显式含有时间 t, 即 $\frac{\partial H}{\partial t}=0$, 则时间 $t=x^{n+1}$ 为循环坐标 (参见 §31.2) , 并且对应的动量 $p_{n+1}=E$ 守恒.

2. 泊松括号

定义 34.2. 在 $2n$ 维相空间中的反称度量设为

$$\begin{pmatrix} O_n & I_n \\ -I_n & O_n \end{pmatrix};$$

其中的两个函数 $f(x,p)$ 和 $g(x,p)$ 的泊松括号是指它们的梯度的内积:

$$\{f,g\} = \langle \nabla f, \nabla g \rangle = g^{ij}\frac{\partial f}{\partial y^i}\frac{\partial g}{\partial y^j} = \sum_{i=1}^{n}\left(\frac{\partial f}{\partial x^i}\frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial x^i}\frac{\partial f}{\partial p_i}\right), \tag{15}$$

其中 $(\nabla f)^i = g^{ij}\dfrac{\partial f}{\partial y^j}$, 等等.

我们有下面的定理:

定理 34.1. 泊松括号有下列性质:

1) $\{f,g\} = -\{g,f\}$, $\{\lambda f_1 + \mu f_2, g\} = \lambda\{f_1,g\} + \mu\{f_2,g\}$,
其中 λ,μ 为常数;

2) $\{f,\{g,h\}\} + \{h,\{f,g\}\} + \{g,\{h,f\}\} = 0$ (16)
(雅可比等式);

3) $\{fg,h\} = f\{g,h\} + g\{f,h\}$; (17)

4) $\nabla\{f,g\} = -[\nabla f, \nabla g]$,
其中 $[,]$ 为向量场的换位子.

证明 性质 1) 和 3) 是显然的.

考虑映射 $f \mapsto \nabla f$: 由相空间中光滑函数的集合到向量场. 其线性性和性质 4) 表明这个映射为代数的同态, 一边是以泊松括号为运算的函数的代数, 另一边是以通常的换位子为运算的向量场代数. 事实上, 这是以相同梯度为等价关系的函数类的代数到向量场的李代数的一个单同态. 因为后面这个代数满足雅可比恒等式, 故第一个代数也满足. 如果我们能证明性质 4), 那么性质 2) 便得证.

在坐标 (x,p) 下我们有

$$\nabla f = \left(\frac{\partial f}{\partial p}, -\frac{\partial f}{\partial x}\right), \quad \nabla g = \left(\frac{\partial g}{\partial p}, -\frac{\partial g}{\partial x}\right).$$

我们来计算这两个向量场的换位子. 我们记得在 §23 中, 向量场 X 和 Y 的换位子在坐标 y^i 下由公式

$$[X,Y]^i = X^j\frac{\partial Y^i}{\partial y^j} - Y^j\frac{\partial X^i}{\partial y^j}.$$

运用此等式于换位子 $[\nabla f, \nabla g]$ 的第一个 x 坐标, 我们有

$$\begin{aligned}[\nabla f, \nabla g]^1 &= \frac{\partial f}{\partial p_j}\frac{\partial}{\partial x^j}\left(\frac{\partial g}{\partial p_1}\right) - \frac{\partial f}{\partial x^j}\frac{\partial}{\partial p_j}\left(\frac{\partial g}{\partial p_1}\right) - \frac{\partial g}{\partial p_j}\frac{\partial}{\partial x^j}\left(\frac{\partial f}{\partial p_1}\right) + \frac{\partial g}{\partial x^j}\frac{\partial}{\partial p_j}\left(\frac{\partial f}{\partial p_1}\right)\\ &= \frac{\partial f}{\partial p_j}\frac{\partial}{\partial p_1}\left(\frac{\partial g}{\partial x^j}\right) + \frac{\partial g}{\partial x^j}\frac{\partial}{\partial p_1}\left(\frac{\partial f}{\partial p_j}\right) - \frac{\partial f}{\partial x^j}\frac{\partial}{\partial p_1}\left(\frac{\partial g}{\partial p_j}\right) - \frac{\partial g}{\partial p_j}\frac{\partial}{\partial p_1}\left(\frac{\partial f}{\partial x^j}\right)\\ &- \frac{\partial}{\partial p_1}\left(\frac{\partial f}{\partial p_j}\frac{\partial g}{\partial x^j} - \frac{\partial f}{\partial x^j}\frac{\partial g}{\partial p_j}\right) = (-\nabla\{f,g\})^1;\end{aligned}$$

在这里我们利用了等式

$$\frac{\partial}{\partial x}\left(\frac{\partial g}{\partial p}\right) = \frac{\partial}{\partial p}\left(\frac{\partial g}{\partial x}\right),\quad \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial p}\right) = \frac{\partial}{\partial p}\left(\frac{\partial f}{\partial x}\right).$$

对于 p 坐标等式 4) 的证明是类似的. 证完. □

推论　相空间上的函数 $f(x,p)$ 在泊松括号下构成了李代数.

我们注意到, 由于定理 34.1 的论断 4), 这个李代数同构于在反称度量 g_{ij} 下梯度向量场的李代数.

现在设 g_{ij} 为任意的反称 (非异) 度量; $\Omega = g_{ij}dy^i \wedge dy^j$. 我们已用前面的公式定义过了泊松括号 $\{f,g\}$:

$$\{f,g\} = g^{ij}\frac{\partial f}{\partial y^i}\frac{\partial g}{\partial y^j}\ (g^{ik}g_{kj} = \delta^i_j). \tag{18}$$

Ω 应满足什么条件才能使光滑函数对于换位子 $\{,\}$ 构成李代数? 下面的定理给出了回答.

定理 34.2.　光滑函数关于泊松括号 (18) 构成李代数当且仅当形式 Ω 为闭, 即 $d\Omega = 0$.

证明　只需要验证雅可比恒等式成立. 它等价于 (参看定理 34.1 的证明) 恒等式

$$\nabla\{f,g\} = -[\nabla f, \nabla g]. \tag{19}$$

等式 (19) 的左端的形式 (第 k 个坐标) 为

$$\nabla\{f,g\}^k = g^{kl}\frac{\partial g^{ij}}{\partial y^l}\frac{\partial f}{\partial y^i}\frac{\partial g}{\partial y^j} + g^{kl}g^{ij}\left(\frac{\partial^2 f}{\partial y^l\partial y^i}\frac{\partial g}{\partial y^j} + \frac{\partial f}{\partial y^i}\frac{\partial^2 g}{\partial y^l\partial y^j}\right);$$

右端为

$$\begin{aligned}-[\nabla f, \nabla g]^k = g^{il}\frac{\partial g^{kj}}{\partial y^i}\frac{\partial g}{\partial y^l}\frac{\partial f}{\partial y^j} - g^{ij}\frac{\partial g^{kl}}{\partial y^i}\frac{\partial f}{\partial y^j}\frac{\partial g}{\partial y^l} +\\ g^{ij}g^{kl}\frac{\partial g}{\partial y^j}\frac{\partial^2 f}{\partial y^l\partial y^i} - g^{ji}g^{kl}\frac{\partial f}{\partial y^i}\frac{\partial^2 g}{\partial y^j\partial y^l}.\end{aligned}$$

使它们相等, 那么在对取和指标缩减和重新编号之后我们得到:(19) 等价于

$$\left(g^{ki}\frac{\partial g^{jl}}{\partial y^i} + g^{li}\frac{\partial g^{kj}}{\partial y^i} + g^{ij}\frac{\partial g^{kl}}{\partial y^i}\right)\frac{\partial f}{\partial y^j}\frac{\partial g}{\partial y^l} = 0.$$

由于 $\frac{\partial f}{\partial y^j}, \frac{\partial g}{\partial y^l}$ 的任意性, 这个等式等价于下面等式:

$$g^{ki}\frac{\partial g^{jl}}{\partial y^i}+g^{li}\frac{\partial g^{kj}}{\partial y^i}+g^{ij}\frac{\partial g^{kl}}{\partial y^i}=0,\quad j,l,k=1,\cdots,2n. \tag{20}$$

对 (20) 乘以 $g_{rk}g_{pj}g_{ql}$ 并利用关系

$$g_{ij}\frac{\partial g^{jk}}{\partial y^l}=-g^{jk}\frac{\partial g_{ij}}{\partial y^l},$$

我们得到

$$\frac{\partial g_{pq}}{\partial y^r}+\frac{\partial g_{qr}}{\partial y^p}+\frac{\partial g_{rp}}{\partial y^q}=0,$$

这就是 Ω 为闭形式的条件. 证完. □

我们现在来解释如下问题: 在什么情形下, 一个 2 阶非异形式 Ω (即 $\Omega^n\neq 0$) 可以经坐标变换化为典则形式 (7)? 因为这样的典则形式是闭的, 故而形式 Ω 为闭的条件是有这种约化的必要条件. 充分性由下面的论断得出, 但我们只加引述而不证明.

定理 34.3. **(达布)** 设 Ω 为 2 阶微分形式, 且 $\Omega^n=\Omega\wedge\cdots\wedge\Omega\neq 0$. 如果形式 Ω 为闭, 则存在局部坐标 $x^1,\cdots,x^n,p_1,\cdots,p_n$, 在此坐标下 Ω 具有典则形式:

$$\Omega=\sum_i dx^i\wedge dp_i.$$

假设我们在相空间中给出了哈密顿 H 的哈密顿系统 (8), 于是沿此系统上任意函数 $f=f(x,p)$ 的导数有形式

$$\dot f=\{f,H\}. \tag{21}$$

特别,$\dot x^i=\{x^i,H\}=\frac{\partial H}{\partial p_i}, \dot p_i=\{p_i,H\}=-\frac{\partial H}{\partial x^i}$.

我们称函数 $f(x,p)$ 为哈密顿系统的积分是说 f 沿哈密顿系统的轨线为常数. 那么, 由定理 34.1 立即得到

结论 如果函数 $f(x,p)$ 与哈密顿 H 交换:$\{f,H\}=0$, 则它是此哈密顿系统的积分 (运动的积分). 哈密顿系统的全部积分函数构成了李代数, 而此代数在函数的乘积下也是封闭的.

我们假定哈密顿 $H(x,p)$ 与时间显式无关:$\frac{\partial H}{\partial t}=0$. 于是能量 $E=H(x,p)$ 守恒, 从而哈密顿 $H(x,p)$ 的哈密顿系统的所有轨线都位于水平曲面 $H(x,p)=E$ 上. 设 $f=f(x,p)$ 为此系统的积分,$\{f,H\}=0$. 但由此也有 $\{H,f\}=0$, 即 $H(x,p)$ 沿哈密顿系统 ∇f 的轨线上为常值. 因此, 向量场 ∇f 切于水平曲面 $H(x,p)=E$. 因为切于已知曲面的向量场是全体向量场的李代数的子代数 (参看 §24.1), 故而谈及在所给水平能量曲面 $H(x,p)=E$ 上哈密顿系统的积分函数的李代数是有意义的.

例 34.1. 设拉格朗日 $L(x,\dot{x})$ 是球面对称的. 正如我们在前面所看到的, 在 $SO(3)$ 中的三个单参数子群为三个积分所对应 (动量矩积分), 其中

$$M_x = L_x^i \frac{\partial L}{\partial \dot{x}^i} = yp_z - zp_y, \quad M_y = L_y^i \frac{\partial L}{\partial \dot{x}^i} = zp_x - xp_z,$$

$$M_z = L_z^i \frac{\partial L}{\partial \dot{x}^i} = xp_y - yp_x, \tag{22}$$

其中 L_x, L_y, L_z 为对应的线性向量场. 换句话说, 函数 M_x, M_y, M_z 为由 L 给出的哈密顿系统的积分 (函数). 向量场 L_x, L_y, L_z 的换位子有形状

$$[L_x, L_y] = L_z, \quad [L_y, L_z] = L_x, \quad [L_z, L_x] = L_y. \tag{23}$$

由定理 34.1 的断言 4) 得出, 动量矩的泊松括号由下面公式计算

$$\{M_x, M_y\} = -M_z, \quad \{M_y, M_z\} = -M_x, \quad \{M_z, M_x\} = -M_y. \tag{24}$$

结论 在球面对称系统的相空间上的函数 M_x, M_y, M_z 关于泊松括号构成了李代数, 它同构于李代数 $so(3)$.

例 34.2. 在开普勒问题中,$\mathbb{R}^6$ 的哈密顿 $H(x,p)$ 的形式为

$$H(x,p) = \frac{p^2}{2m} + \frac{\alpha}{|x|}, \quad \alpha < 0. \tag{25}$$

由于球面对称性, 这里有三个动量矩的积分函数

$$M = (M_1, M_2, M_3) = [x, p] = 常数. \tag{26}$$

原来, 在此问题中还有三个积分

$$W = (W_1, W_2, W_3) = \left[\frac{p}{m}, M\right] + \frac{\alpha x}{|x|} = 常数 \tag{27}$$

(拉普拉斯 – 龙格 – 楞茨向量).我们来计算泊松括号 $\{M_i, W_i\}, \{W_i, W_j\}$. 对此, 我们要运用下列对函数 p_i, M_j 的泊松表达式:

$$\begin{gathered}\{p_i, M_i\} = 0, \quad \{p_1, M_2\} = p_3, \quad \{p_1, M_3\} = -p_2, \quad \{p_2, M_1\} = -p_3, \\ \{p_2, M_3\} = p_1, \quad \{p_3, M_1\} = p_2, \quad \{p_3, M_2\} = -p_1\end{gathered} \tag{28}$$

(请验证!).(设 $x^i = x_i, i = 1, 2, 3$) 我们有

$$M_1 = x_2 p_3 - x_3 p_2,$$
$$W_1 = \frac{1}{m}(p_2 M_3 - p_3 M_2) + \frac{\alpha x_1}{|x|}, \quad W_2 = \frac{1}{m}(p_3 M_1 - p_1 M_3) + \frac{\alpha x_2}{|x|}.$$

利用定理 34.1 的性质 1),3) 和公式 (24),(28), 我们得到

$$\begin{aligned}\{M_1, W_1\} &= \frac{1}{m}\{M_1, p_2M_3 - p_3M_2\} + \alpha\{M_1, \frac{x_1}{|x|}\}\\ &= \frac{1}{m}(\{M_1, p_2\}M_3 + \{M_1, M_3\}p_2 - \{M_1, p_3\}M_2 - \{M_1, M_2\}p_3) +\\ &\quad \frac{\alpha}{|x|}\{M_1, x_1\} + \alpha x_1\{M_1, \frac{1}{|x|}\}\\ &= \frac{1}{m}(p_3M_3 - M_2p_2 + p_2M_2 - p_3M_3) = 0,\end{aligned}$$

$$\begin{aligned}\{M_1, W_2\} &= \frac{1}{m}\{M_1, p_3M_1 - p_1M_3\} + \alpha\{M_1, \frac{x_2}{|x|}\}\\ &= \frac{1}{m}(\{M_1, p_3\}M_1 - \{M_1, M_3\}p_1) + \frac{\alpha}{|x|}\{M_1, x_2\}\\ &= \frac{1}{m}(-p_2M_1 + p_1M_2) + \frac{\alpha x_3}{|x|} = W_3.\end{aligned}$$

按同样的方式计算余下的括号 $\{M_i, W_j\}$. 我们得到下列泊松括号的表格:

$\{M_i, W_j\}$	W_1	W_2	W_3
M_1	0	W_3	$-W_2$
M_2	$-W_3$	0	W_1
M_3	W_2	$-W_1$	0

(29)

类似的, 然而更加长的计算能够在常值水平能量 $H(x,p) = E$ 的条件下求出形如 $\{W_i, W_j\}$ 的两两括号. 我们有

$$\{W_1, W_2\} = -\frac{2E}{m}M_3, \quad \{W_2, W_3\} = -\frac{2E}{m}M_1,$$

$$\{W_3, W_1\} = -\frac{2E}{m}M_2. \tag{30}$$

显见, 在常值水平能量 $H = E$ 条件下, 函数 W_i, M_j 的李代数结构实质上依赖于这个水平能量 (参看下面的习题 34.3).

3. 典则变换

设有哈密顿函数 $H(x,p)$ 的任意哈密顿系统:

$$\dot{x}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial x^i}, \quad i = 1, \cdots, n.$$

定理 34.4. 对于哈密顿系统, 形式 $\Omega = \sum_{i=1}^{n} dx^i \wedge dp_i$ 保持不变, 即对于任意哈密顿 $H = H(x,p)$, 沿向量场 ∇H 形式 Ω 的李导数等于零

$$\dot{\Omega} = L_{\nabla H}\Omega = 0. \tag{31}$$

证明 为了计算形式 Ω 沿向量场 ∇H 的导数 (即李导数 $\dot{\Omega} = L_{\nabla H}\Omega$), 需要利用下面的事实 (见 §23.2)

$$(\Omega_1 \wedge \Omega_2)^{\cdot} = \dot{\Omega}_1 \wedge \Omega_2 + \Omega_1 \wedge \dot{\Omega}_2, \tag{32}$$

$$(dx^i)^{\cdot} = d\left(\frac{\partial H}{\partial p_i}\right) = \frac{\partial^2 H}{\partial p_i \partial x^j} dx^j + \frac{\partial^2 H}{\partial p_i \partial p_j} dp_j, \tag{33}$$

$$(dp_i)^{\cdot} = -d\left(\frac{\partial H}{\partial x^i}\right) = -\frac{\partial^2 H}{\partial x^i \partial x^j} dx^j - \frac{\partial^2 H}{\partial x^i \partial p_j} dp_j. \tag{34}$$

故而

$$\begin{aligned}\left(\sum_i dx^i \wedge dp_i\right)^{\cdot} &= \sum_{i,j}\left\{\frac{\partial^2 H}{\partial p_i \partial x^j} dx^j \wedge dp_i + \frac{\partial^2 H}{\partial p_i \partial p_j} dp_j \wedge dp_i - \right. \\ &\qquad \left. dx^i \wedge \frac{\partial^2 H}{\partial x^i \partial x^j} dx^j - dx^i \wedge \frac{\partial^2 H}{\partial x^i \partial p_j} dp_j\right\} \\ &= 0,\end{aligned}$$

这是因为外积是反称的, 故有此结果. □

推论 (刘维尔) (x,p) 空间中任意区域的相空间体积对于哈密顿系统保持不变.

证明 因为 $\dot{\Omega} = 0$, 故由 (32) 得到 $(\Omega \wedge \cdots \wedge \Omega)^{\cdot} = 0$. 根据引理 34.2, 在相空间中的体积元恰好是形式 Ω 自己的 n 重外积. 引理得证. □

定义 34.3. 称相空间 (x,p) 到自身的变换 $\Phi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ 为典则变换是说, 这个哈密顿系统保持空间的反称内积不变 (即将 Ω 变到 Ω).

换句话说, 典则变换就是反称度量下的运动. 由定理 34.4 知, 沿哈密顿 $H(x,p)$ 的哈密顿系统的轨线的位移给出了典则变换的单参数群. 反之也对. 设给出了典则变换的任意局部单参数群 $\Phi_t(x,p) = (x(t),p(t))$, 并设 $X = \frac{d\Phi_t}{dt}\big|_{t=0}$ 为对应的向量场. 则有

定理 34.5. 存在局部唯一的光滑哈密顿函数 $H(x,p)$, 使得向量场 X 对于它是哈密顿的, 即有

$$X = \left(\frac{\partial H}{\partial p}, -\frac{\partial H}{\partial x}\right).$$

证明 设在坐标 (x,p) 下, 场 X 为 $X=(A^i,B_i), i=1,\cdots,n$. 考虑时间位移 Δt. 我们有

$$\Phi_{\Delta t}:\begin{cases} x^i\to x^i+A^i(x,p)\Delta t+O(\Delta t^2)=x'^i,\\ p_i\to p_i+B_i(x,p)\Delta t+O(\Delta t^2)=p_i'.\end{cases}$$

定理的条件要求保持反称度量不变

$$\sum_i dx'^i\wedge dp_i'=\sum_i dx^i\wedge dp_i.$$

因而由

$$\begin{aligned} dx'^i\wedge dp_i' &=(dx^i+(dA^i)\Delta t)\wedge(dp_i+(dB_i)\Delta t)+O(\Delta t^2)\\ &=dx^i\wedge dp_i+\Delta t\left[\frac{\partial A^i}{\partial x^j}dx^j\wedge dp_i+\frac{\partial A^i}{\partial p_j}dp_j\wedge dp_i+\frac{\partial B_i}{\partial p_j}dx^i\wedge dp_j+\right.\\ &\qquad\left. x\frac{\partial B_i}{\partial x^j}dx^i\wedge dx^j\right]+O(\Delta t^2)\end{aligned}$$

得到

1) $\dfrac{\partial A^i}{\partial p_j}dp_i\wedge dp_j=0$ 即 $\dfrac{\partial A^i}{\partial p_j}=\dfrac{\partial A^j}{\partial p_i}$;

2) $\dfrac{\partial B_i}{\partial x^j}dx^i\wedge dx^j=0$ 即 $\dfrac{\partial B_i}{\partial x^j}=\dfrac{\partial B_j}{\partial x^i}$;

3) $\dfrac{\partial A^i}{\partial x^j}dx^j\wedge dp_i=-\dfrac{\partial B_i}{\partial p_j}dx^i\wedge dp_j$ 即 $\dfrac{\partial A^i}{\partial x^j}=-\dfrac{\partial B_j}{\partial p_i}$.

如果 $A^i=\dfrac{\partial H}{\partial p_i}, B_i=-\dfrac{\partial H}{\partial x^i}$, 则上面这些条件成立. 反之, 由这些条件明显看出, 形式 $A^idp_i-B_idx^i$ 为全微分 (由格林定理), 从而可令 $H(x,p)=\displaystyle\int_{(x_0,p_0)}^{(x,p)}(-B_idx^i+A^idp_i)$. 定理得证. □

注 哈密顿 $H(x,p,t)$ 对时间的显式依赖性给出了典则变换的单参数族 Φ, 然而并不是一个群 $(\Phi_{t_1+t_2}\neq\Phi_{t_1}\circ\Phi_{t_2})$.

当 $n=1$ 时, 形式 $\Omega=dx\wedge dp$ 的守恒性相当于 (x,p) 平面的面积的守恒. 因此此时典则变换类化为使面积不变的变换类. 当 $n>1$ 时, 典则变换类小于使体积不变的类 (可以证明形式 $\Omega\wedge\cdots\wedge\Omega$ 不变时 Ω 却不是不变的).

称相空间 (x,p) 的 $\mathbb{R}^{2n}$ 中的线性典则变换为辛变换 (见定义 14.1). 在 $n=1$ 时全部辛变换的群同于 $SL(2,\mathbb{R})$.

我们来研究辛变换群的李代数. 这个代数中的矩阵为 $K=\dfrac{d}{dt}A(t)\Big|_{t=0}$, 其中 $A(t)$ 是某个光滑的辛变换族, 使 $A(0)=I_{2n}$, 按照定理 34.5, 存在哈密顿 $H(x,p)$ 使得

$$Ky=\left(\frac{\partial H}{\partial p},-\frac{\partial H}{\partial x}\right),$$

其中 $y=(x,p)$. 不需太复杂便能知道, 在此情形下,$H(x,p)$ 是个二次型

$$H=\frac{1}{2}\sum_{i,j=1}^{n}(a_{ij}x^ix^j+b_{ij}x^ip_j+c_{ij}p_ip_j). \tag{35}$$

对应于这个二次型的对称矩阵为 $\frac{1}{2}\begin{pmatrix} A & B \\ B^{\mathrm{T}} & C \end{pmatrix}$, $A=(a_{ij}), B=(b_{ij}), C=(c_{ij})$, 其中矩阵 A 和 C 为对称. 因而辛变换群的李代数中的矩阵 K 的形式为

$$K=\begin{pmatrix} B^{\mathrm{T}} & C \\ -A & -B \end{pmatrix},\quad A=A^{\mathrm{T}},\quad C-C^{\mathrm{T}}. \tag{36}$$

于是, 辛变换群的李代数同构于二次哈密顿 (35) 在泊松括号下的李代数. 简谐振动给出了二次哈密顿的简单例子, 其中的 $H=\frac{p^2}{2m}+\frac{m\omega^2x^2}{2}$,$\omega$ 为频率. 因此也称二次哈密顿为 "广义振子".

习题

34.1. 对构形空间 (即具坐标 x 的空间) 的向量场 X 给予相空间中的函数 F_X, 其中 $F_X=p_iX^i$. 证明 $\{F_X,F_Y\}=-F_{[X,Y]}$.

34.2. 设在相空间上的函数 $f=f(x)$ 与动量 p 无关. 证明

$$\{f,F_X\}=\partial_X f.$$

34.3. 证明在开普勒问题中, 在固定能量 E 下的积分函数 W_i, M_j 的李代数同构于:

a) 当 $E<0$ 时,$so(4)$;

b) 当 $E=0$ 时,$\mathbb{R}^3$ 中运动群的李代数;

c) 当 $E>0$ 时,$so(1,3)$.

34.4. 设 $\Omega=g_{ik}dy^i\wedge dy^k$ 为辛矩阵,$X=(X^k)$ 为向量场. 验证 $g_{ik}X^kdy^i$ 为闭形式当且仅当形式 Ω 沿向量场 X 的李导数为零:$L_X\Omega=0$.

34.5. 设 M_x,M_y,M_z 为在第 2 小节中定义的动量矩积分,$M^2=M_x^2+M_y^2+M_z^2$. 验证

$$\{M^2,M_x\}=\{M^2,M_y\}=\{M^2,M_z\}=0.$$

34.6. 设 $M=(M_1,M_2,M_3)=[x,p]$. 证明函数 p_i,M_j 构成了在泊松括号下的李代数, 这个李代数同构于三维欧氏空间运动群的李代数.(参看习题 11.24)

34.7. 设 $(g_{ij})=-(g_{ji})$ 为非异反称矩阵, 它给出了 $\mathbb{R}^{2n}$ 中的反称内积. 证明 $\mathbb{R}^{2n}$ 中使此反称内积限制在其上为零的线性子空间的维数不大于 n.

34.8. 设在具典则坐标 $x^1,\cdots,x^n,p_1,\cdots p_n$ 的相空间中引进了新坐标 $X^1,\cdots,X^n,P_1,\cdots,P_n$, 使得 $p_i=\dfrac{\partial S(x,X)}{\partial x^i},P_i=-\dfrac{\partial S(x,X)}{\partial X^i},i=1,\cdots,n$, 而 $S(x,X)$ 为变量 x^i,X^i 的某个光滑函数 (设 $\det\left(\dfrac{\partial^2 S}{\partial x^i\partial X^j}\right)\neq 0$). 证明这个新坐标也是典则的 (称函数 $S(x,X)$ 为典则变换 $(x,p)\mapsto(X,P)$ 的生成函数.

§35. 曲面的拉格朗日函数

1. 轨线把和哈密顿 – 雅可比方程

在各种目的中实质上要了解的不仅是哈密顿个别轨线的性质, 而是整个轨线把的性质, 更准确地, 这个问题的提法是: 考虑 $t=0$ 时在 $2n$ 维相空间 (x,p) 中的 n 维曲面 $\Gamma=\Gamma^n$, 其图像为

$$p_i=f_i(x^1,\cdots,x^n),\quad i=1,\cdots,n. \tag{1}$$

曲面 Γ 中的每个点沿具哈密顿 $H(x,p,t)$ 的系统的轨线移动; 在时刻 $t>0$ 得到了位移曲面 Γ_t, 而 $\Gamma_0=\Gamma$. 将曲面的集合 Γ_t 看为扩张了的相空间 (x,p,t,E) 中的 $n+1$ 维曲面比较方便; 点 $(x(t),p(t),E(t),t),E=H(x,p,t)$ 给出了曲面 Γ^{n+1}, 而在 $t=0$ 的截面为 Γ, 并且它由从 Γ 出发的全部轨线构成. 但是, 我们需要某种更广的相似空间.

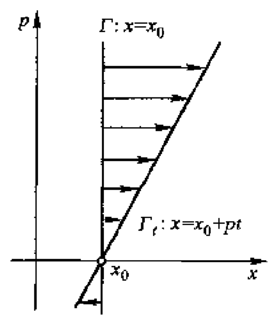


图 35

例如, 设 $\Gamma=\Gamma_0$ 为 $x=x_0$ 的曲面 (动量为任意; 图 35); 我们得到作为由所有 $x=x_0$ 上点出发的轨线把的 Γ^{n+1}. 在 $t=0$ 时曲面 Γ_0 不是作为函数 $p=f(x)$ 的图像, 而 $t>0$ 时则可以是这样的图像 (虽然是局部的).

定义 35.1. 称在相空间 (x,p) 中的曲面 Γ^n 为拉格朗日曲面是说对曲面 Γ 上由任意点 $Q\in\Gamma^n$ 出发的所有曲线 γ,(截尾) 作用 $S=\displaystyle\int_\gamma pdx$ 是只依赖于路径 γ 的终点的局部单值函数.

定义 35.2. 在扩张相空间中的曲面 Γ^{n+1} 被称做拉格朗日曲面是说对 Γ^{n+1} 上所有从点 Q 为起点的路径 γ, 作用 $S=\displaystyle\int_\gamma(pdx-Edt)$(其中 E 为曲面 Γ^{n+1} 上的 $H(x,p,t)$) 是终点的局部单值函数.

引理 35.1. 曲面 Γ^n(分别地,Γ^{n+1}) 如果为图像 $p_i=f_i(x),i=1,\cdots,n$ 的形式 (分别地对 Γ^{n+1} 为 $i=0,1,\cdots,n,p_0=E,x^0=t$), 则它为拉格朗日的当且仅当

$$\frac{\partial S(x)}{\partial x^i}=p_i\quad(\text{对}\Gamma^n) \tag{2}$$

$$\frac{\partial S(x,t)}{\partial x^i}=p_i \text{ 和 } \frac{\partial S(x,t)}{\partial t}=-H(x,p,t) \tag{3}$$

或者

$$\frac{\partial S(x,t)}{\partial t}=-H\left(x,\frac{\partial S}{\partial x},t\right) \quad (\text{对于}\Gamma^{n+1}).$$

证明　设 S 为一个局部单值函数. 因为按定义有 $S=\int pdx-H(x,p,t)dt$, 于是得 $dS=pdx-Hdt$. 由此得到 $p=\frac{\partial S}{\partial x},-H=\frac{\partial S}{\partial t}$. 反之, 如果 $p=\frac{\partial S}{\partial x},-H=\frac{\partial S}{\partial t}$, 则由格林定理知 S 为局部单值. 证完. □

我们称作用 $S(x,t)$ 为轨线把的作用, 而称方程

$$\frac{\partial S}{\partial t}+H\left(x,\frac{\partial S}{\partial x},t\right)=0 \tag{4}$$

为哈密顿－雅可比方程.

如果已知函数 $S(x,t)$, 则曲面 Γ^{n+1} 由方程 $p=\frac{\partial S}{\partial x},E=-\frac{\partial S}{\partial t}$ 给出, 而它在 $t=t_0$ 的截面 $\Gamma_{t_0}^n$ 由从 $\Gamma=\Gamma_0$ 沿哈密顿系统 $\dot{p}=-\frac{\partial H}{\partial x},\dot{x}=\frac{\partial H}{\partial p}$ 的轨线按时间运动得到.

我们现在给出在相空间 (x,p) 中 (或在推广的相空间 (x,p,E,t) 中) 的拉格朗日曲面的另一个定义

定义 35.3.　称在相空间中 n 维曲面 Γ 为拉格朗日曲面是说它的任意切向量的反称内积等于零 (即形式 $\Omega=\sum_i dx^i\wedge dp_i$ 在 Γ 上的限制为零).

这个拉格朗日曲面的定义为几何不变, 并且并不要求它由光滑函数 $p_i=f_i(x^1,\cdots,x^n)$ 表示.

最简单的拉格朗日曲面的形式为

a) $\Gamma_{x_0}=\{x=x_0$, 动量为任意$\}$;

b) $\Gamma_{p_0}=\{p=p_0$, 坐标为任意$\}$.

显然, 在典则变换下, 曲面的拉格朗日仍变为拉格朗日. 譬如变换 $x\to p,p\to -x$ 有

$$\Omega=dx^i\wedge dp_i\to -dp_i\wedge dx^i=dx^i\wedge dp_i=\Omega.$$

这时曲面 Γ_{x_0} 和 Γ_{p_0} 互换.

定理 35.1.　如果由图像 $p_i=f_i(x)$ 给出的曲面在定义 35.3 的意义下是拉格朗日的, 则存在函数 $S(x)$ 使得 $f_i=\frac{\partial S(x)}{\partial x^i}$, 反之亦然.

证明 如果曲面 Γ 由方程 $p_i = f_i(x)$ 给出, 则形式 Ω 在 Γ 上的限制为

$$\Omega|_\Gamma = \sum_i dx^i \wedge dp_i(x) = \sum_{i,j} dx^i \wedge \frac{\partial f_i}{\partial x^j} dx^j.$$

因此 $\Omega|_\Gamma = \sum_{i<j} \left(\frac{\partial f_i}{\partial x^j} - \frac{\partial f_j}{\partial x^i}\right) dx^i \wedge dx^j$. 如果 $\Omega|_\Gamma \equiv 0$, 则 $\frac{\partial f_i}{\partial x^j} \equiv \frac{\partial f_j}{\partial x^i}$. 这个条件是使积分 $S(x) = \int_{x_0}^{x} f_i(x) dx^i$ 给出的函数不依于路径的充要条件 (格林公式). 从而得到了我们的断言. □

显然, 对于扩张的相空间 (x, p, E, t) 成立有类似的命题.

设 $H(x,p) = H$ 不显式依赖于 t. 于是有

定理 35.2.

a) 向量 $\nabla H = \left(\frac{\partial H}{\partial p}, -\frac{\partial H}{\partial x}\right) = (\dot{x}, \dot{p})$ 切于曲面 $H = E_0 =$ 常数;

b) 向量 ∇H 与所有切于水平能量 $H = E_0$ 的内积为零;

c) 任意位于水平能量 $H(x,p) = E_0$ 上的 n 维拉格朗日曲面在其自身的每点上向量 ∇H 都是它的切向量. 特别, 当具哈密顿 H 的哈密顿系统的轨线与此曲面至少有一个公共点时, 这个曲面就包含了整个这条轨线.

证明 向量 ξ 切于曲面 $H = E_0$ 的充要条件由方程 $\xi^i \frac{\partial H}{\partial y^i} = 0, y = (x,p)$ 给出. 由此立刻得到 a). 另外

$$\langle \xi, \nabla H \rangle = \xi^i g_{ij} (\nabla H^j) = \xi^i g_{ij} g^{jk} \frac{\partial H}{\partial y^k} = \xi^i \frac{\partial H}{\partial y^i} = 0.$$

于是论断 b) 得证, 我们来证明论断 c). 设 P 为拉格朗日曲面 Γ(维数为 n) 的点, Γ 整个位于水平能量 $H(x,p) = E_0$ 上, 并设 $\xi_1, \cdots, \xi_n$ 为曲面 Γ 在点 P 的切空间的基. 按照拉格朗日的性质, 我们有 $\langle \xi_i, \xi_k \rangle = 0$. 考虑在点 P 的向量 ∇H. 由 b) 得到 $\langle \nabla H, \xi_i \rangle = 0, i = 1, \cdots, n$. 内积 $\langle , \rangle$ 非异 $(\det(g_{ij}) \neq 0)$. 故而 $\nabla H = \sum_i \lambda_i \xi_i$(我们记得, 内积为零的线性子空间的维数不大于 n. 参看习题 34.7). 因此,∇H 切于 Γ. 这对曲面 Γ 的所有点都成立. 所以在 Γ 的每个点上引出的哈密顿 H 的轨线都切于它 (从而表明位于其上). 定理得证. □

推论 考虑在水平面 $H(x,p) = E_0$ 上的任意 $n-1$ 维拉格朗日曲面 S^{n-1}. 由它的每点我们引出在 (x,p) 空间中哈密顿系统的轨线. 我们假定这样得到的曲面是 n 维的. 于是这个 n 维曲面 Γ^n 便是拉格朗日的, 并且位于水平能量 E_0 上. 如果 (局部地) 曲面 Γ^n 由方程

$$p_i = f_i(x) = \frac{\partial S_0}{\partial x^i} \tag{5}$$

给出，那么函数 $S_0(x)$ 满足哈密顿 – 雅可比截尾方程

$$E_0 = H\left(x, \frac{\partial S_0}{\partial x}\right). \tag{6}$$

这时的函数 $S(x,t) = \int p dx - H dt$ 有形式

$$S(x,t) = -E_0 t + S_0(x), \quad -\frac{\partial S}{\partial t} = H\left(x, \frac{\partial S}{\partial x}\right).$$

2. 作为动量的一阶齐次函数的哈密顿情形

对于动量为一阶齐次的哈密顿函数 $H(x,p)$

$$H(x,\lambda p) = \lambda H(x,p), \quad \lambda > 0, \tag{7}$$

我们应该单独地考虑这种特殊的哈密顿体系. 例如, 对在各向同性的介质中的光线, 我们有哈密顿 $H(x,p) = c(x)|p|$. 这样的哈密顿只在 $p \neq 0$ 有定义, 并且只需了解在一个水平能量 $H = E_0$ (例如 $E_0 = 1$) 时的轨线就够了. 其余的轨线可以通过相似变换 $p \to \lambda p, H \to \lambda H(x \to x)$ 得到.

我们也注意到, 通常由拉格朗日 $L = g_{ij}v^i v^j$ 得出在度量 $g_{ij}(x)$ 下的测地线, 而这个度量对应的哈密顿为 $H = g^{ij}p_i p_j$, 但是它们也可由哈密顿 $H' = \sqrt{H} = \sqrt{g^{ij}p_i p_j}$ 得到 (在水平能量 $H =$ 常数上, 对应的哈密顿方程与原来方程以常数因子成比例). 同时也有 $H'(x,\lambda p) = \lambda H'(x,p), \lambda > 0$.

定理 35.3. 如果 (局部的) 单参数变换群 Φ_t 由哈密顿 $H(x,p)$ 得到, 而 $H(x,p)$ 满足 $H(x,\lambda p) = \lambda H(x,p), \lambda > 0$, 于是所有变换 Φ_t 保持形式 $pdx = p_i dx^i$ 不变, 即余向量 $(p,0)$ 沿向量场 $X = (\dot{x},\dot{p}) = \left(\frac{\partial H}{\partial p}, -\frac{\partial H}{\partial x}\right)$ 的李导数等于零.

证明 考虑对于小 Δt 的变换 $\Phi_{\Delta t}$. 我们有

$$p_i \to p_i' = p_i - \frac{\partial H}{\partial x^i}\Delta t_i, \quad x^i \to x'^i = x^i + \frac{\partial H}{\partial p_i}\Delta t,$$

$$dx^i \to dx'^i + \Delta t d\left(\frac{\partial H}{\partial p_i}\right).$$

这些等式在 $O(\Delta t^2)$ 精度下成立, 这是因为 (在同样的精度下)

$$\Phi_{\Delta t}(p_i) = p_i' = p_i + \dot{p}_i \Delta t,$$

$$\Phi_{\Delta t}(x^i) = x'^i = x^i + \dot{x}^i \Delta t.$$

对于形式 pdx 有

$$\begin{aligned} p_i' dx'^i &= \left(p_i - \Delta t \frac{\partial H}{\partial x^i}\right)\left(dx^i + \Delta t d\left(\frac{\partial H}{\partial p_i}\right)\right) \\ &= p_i dx^i + \Delta t\left[-\frac{\partial H}{\partial x^i}dx^i + p_i d\left(\frac{\partial H}{\partial p_i}\right)\right] + O(\Delta t^2). \end{aligned}$$

因为 $p_i d\left(\frac{\partial H}{\partial p_i}\right) = d\left(p_i \frac{\partial H}{\partial p_i}\right) - \frac{\partial H}{\partial p_i} dp_i$, 故而

$$p_i' dx'^i = p_i dx^i + \Delta t\left[-dH + d\left(p_i \frac{\partial H}{\partial p_i}\right)\right] + O(\Delta t^2)$$

或者

$$p' dx' = p dx + \Delta t df + O(\Delta t^2),$$

其中 $f = p\frac{\partial H}{\partial p} - H = L$. 如果 $H(x, \lambda p) = \lambda H(x, p)$, 则

$$p\frac{\partial H}{\partial p} = \frac{\partial H(x, \lambda p)}{\partial \lambda} = H(x, p).$$

因此,$p'dx' = pdx + O(\Delta t^2)$; 这表明李导数 $\frac{d}{dt}(pdx)$ 等于零, 即形式 pdx 相对于时间不变. 以此可验证并直接计算余向量场 $T = (T_j) = (p_1, \cdots, p_n, 0, \cdots, 0)$ 沿向量 $X = \left(\frac{\partial H}{\partial p}, -\frac{\partial H}{\partial x}\right)$ 的李导数: $(L_X T)_j = X^k \frac{\partial T_j}{\partial y^k} + T_k \frac{\partial X^k}{\partial p^j}, y = (x, p)$. 当 $j = 1, \cdots, n$ 时我们有

$$\begin{aligned}(L_X T)_j &= X^k \frac{\partial p_j}{\partial y^k} + p_l \frac{\partial X^l}{\partial x^j} = X^{j+n} + p_l \frac{\partial^2 H}{\partial x^j \partial p_l} \\ &= -\frac{\partial H}{\partial x^j} + \frac{\partial}{\partial x^j}\left(p_l \frac{\partial H}{\partial p_l}\right) = 0.\end{aligned}$$

类似地, 对于 $j = n+1, \cdots, 2n$ 有

$$(L_X T)_j = p_l \frac{\partial X^l}{\partial p_j} = \frac{\partial}{\partial p_j}\left(p_l \frac{\partial H}{\partial p_l}\right) - \frac{\partial H}{\partial p_j} = 0.$$

定理证毕. □

定义 35.4. 称相空间中曲面 Γ 为拉格朗日圆锥曲面是说形式 pdx 在曲面 Γ 上的限制恒等于零.

由定理 35.3 得到

推论 满足 $H(x, \lambda p) = \lambda H(x, p)$ 的哈密顿系统保持拉格朗日圆锥曲面类不变.

曲面 $\Gamma_{x_0}(x = x_0$, 动量为任意) 是圆锥的, 这是因为在其上 $pdx = 0$. 另一方面 $\Gamma_{p_0}(p = p_0 \neq 0$, 坐标为任意) 则不是圆锥的 $(pdx \neq 0)$.

对任意满足关系 $H(x, \lambda p) = \lambda H(x, p)$ 的哈密顿, 其曲面 Γ_{x_0} 与从一个点出发的重要轨线把有关. 考虑在 Γ_{x_0} 内的 $(n-1)$ 维曲面 $H = E_0$; 我们以 $S_{x_0}^{n-1}$ 表示这个 $n-1$ 维的曲面. 考虑所有起点只在 $S_{x_0}^{n-1}$ 中的轨线. 在任意时刻 $t > 0$ 我们得到了曲面 $S_{x_0}^{n-1}(t)$.

定义 35.5. 称相空间的曲面 $S_{x_0}^{n-1}(t)$ 在 x 空间上的投影为在时刻 $t>0$ 时对于由点 x_0, 当 $t=t_0$ 时出发的轨线把的波前 (这里的 $S_{x_0}^{n-1}(t)$ 的投影是在 x 空间中的 $n-1$ 维曲面, 它依赖于时间 t).

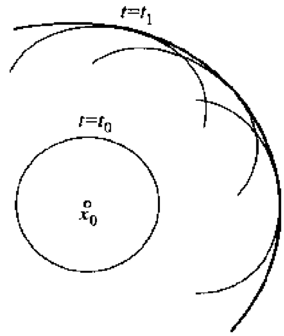


图 36

显然, 在时刻 $t=0$ 曲面 $S_{x_0}^{n-1}(0)$ 在 x 空间的投影是一个点 x_0(由定义).

对于波前曲面成立 "惠更斯原理", 它断言在 $t_1>t_0>0$ 时可以由在时刻 t_0 的波前面得到在时刻 t_1 的波前面. 原来, 为此需要取在 $t=t_0$ 时的波前, 然后以它的所有的点为中心 (或起点) 需要再考虑对应于时间 t_1-t_0 的波前曲面. 包括所有这些曲面的曲面便是在时刻 t_1 的波前 (图 36).

现在设有一个任意的拉格朗日圆锥曲面 Γ^n, 即使得 $pdx|_{\Gamma^n}=0$ 成立的曲面.

引理 35.2. 拉格朗日圆锥面 Γ^n 在 x 空间中的投影的维数 $\leqslant n-1$.

证明 Γ 上的切向量的 x 分量与线性关系 $pdx=0$ 相关联. 所以在投影上切空间的维数不超过 $n-1$. 引理得证. □

因此, 曲面 Γ^n 甚至不能局部表示为图像 $p_i=f_i(x)$ 的形式, 也不与函数 $S(x)$ 有所关联. 故而我们像对所分析的曲面 Γ_{x_0} 的例子那样, 在 Γ^n 中挑出 $H(x,p)=E_0$ 的那一部分, 记之为 S^{n-1}. 根据引理 35.2, 我们从 S^{n-1} 引出轨线把, 并得到拉格朗日曲面 $\widetilde{\Gamma}^n=\bigcup\limits_{-\infty<t<\infty}S^{n-1}(t)$, 它位于水平曲面 $H(x,p)=E_0$ 中, 整体由轨线组成. 曲面 $\widetilde{\Gamma}^n$ 已局部地有形状 $p_i=\dfrac{\partial S_0(x)}{\partial x^i}$. 对于函数 $S_0(x)$ 我们则有哈密顿 – 雅可比截尾方程

$$E_0=H\left(x,\frac{\partial S_0(x)}{\partial x}\right). \tag{8}$$

如果哈密顿不显式依赖于时间, 则可以做成这个运算.

因为曲面 Γ^n 是拉格朗日圆锥的, 故形式 pdx 在 S^{n-1} 等于零. 因此这个形式也在每个曲面 $S^{n-1}(t)$ 上等于零 (由定理 35.3). 故给出拉格朗日曲面 $\widetilde{\Gamma}^n$ 的函数 $S_0(P)=\displaystyle\int_{P_0}^{P}pdx$ 在曲面 $S^{n-1}(t)$ 上为常数, 其中 t 为任意. 另外,$\widetilde{\Gamma}^n$ 位于水平能量 $H=E_0$ 上. 故 $E_0=H\left(x,\dfrac{\partial S_0}{\partial x}\right)$.

这样, 水平曲面 $S_0(x)$= 常数恰好与曲面 $S^{n-1}(t)$ 在 x 空间的投影重合. 于是我们已经证明了

定理 35.4. 水平曲面 $S_0(x)=$ 常数同于波前曲面 (即同于 $S^{n-1}(t)$ 在 x 空间中的投影), 其中 S_0 为上面所定义的截尾作用函数 S_0.

习题

35.1. 设在相空间 $\mathbb{R}^{2n}$ 中给出了 n 个独立的函数 $f_1(x,p),\cdots,f_n(x,p)$, 它们两两交换:$\{f_i,f_j\}=0$. 证明在 $\mathbb{R}^{2n}$ 中由方程组 $f_1=0,\cdots,f_n=0$ 所给出的曲面是拉格朗日的.

§36. 测地方程的二阶变分

1. 二阶变分公式

我们在 §31 中曾导出了欧拉 – 拉格朗日方程

$$\frac{\delta S}{\delta x^i}=\frac{\partial L}{\partial x^i}-\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right)=0,\quad i=1,\cdots,n, \tag{1}$$

这是从泛函

$$S[\gamma]=\int\limits_{\gamma} L(x,\dot{x})dt \tag{2}$$

的极值问题得到的; 在起点为 P 终点为 Q 的所有曲线中泛函 $S[\gamma]$ 在某条曲线 γ 具有极小值的必要条件是 γ 满足这个方程. 如果曲线 $\gamma:\{x^i=x^i(t)\}$ 满足欧拉 – 拉格朗日方程, 那么如何求出使此曲线让 $S[\gamma]$ 能真正取得极小的条件?

像我们所知道的, 多元函数 $f(x^1,\cdots,x^N)$ 使得点 P 为极小值点的必要条件是

$$\left.\frac{\partial f}{\partial x^i}\right|_P=0,\quad i=1,\cdots,N,$$

而在满足上述条件下的充分条件是二次型 $\dfrac{\partial^2 f}{\partial x^i\partial x^j}dx^i dx^j$ 在点 P 为正定. 因此为了求出 $S[\gamma]$ 的极小值, 其中 γ 满足欧拉 – 拉格朗日方程, 必须计算类似的双线性形式的二阶微分 (二阶变分)

$$\left.\frac{\partial^2}{\partial\lambda\partial\mu}S[\gamma+\lambda\xi+\mu\eta]\right|_{\substack{\lambda=0\\ \mu=0}}=G_\gamma(\xi,\eta)=G_\gamma(\eta,\xi), \tag{3}$$

其中 η,ξ 为向量场, 它在曲线 $\gamma(t)$ 上给出, 并在点 $\gamma(a)=P$ 和 $\gamma(b)=Q$ 为零.

引理 36.1. 如果 $\gamma:\{x^i=x^i(t)\}$ 满足欧拉 – 拉格朗日方程, 则有公式

$$G_\gamma(\xi,\eta)=\left.\frac{\partial^2 S[\gamma+\lambda\xi+\mu\eta]}{\partial\lambda\partial\mu}\right|_{\substack{\lambda=0\\ \mu=0}}=-\int_a^b (J_{ij}\xi^j)\eta^i dt, \tag{4}$$

其中

$$J_{ij}\xi^j=\frac{d}{dt}\left(\frac{\partial^2 L}{\partial\dot{x}^i\partial\dot{x}^j}\dot{\xi}^j+\frac{\partial^2 L}{\partial\dot{x}^i\partial x^j}\xi^j\right)-\frac{\partial^2 L}{\partial x^i\partial\dot{x}^j}\dot{\xi}^j-\frac{\partial^2 L}{\partial x^i\partial x^j}\xi^j. \tag{5}$$

证明 利用一阶变分公式 (参看 §31) 我们有

$$\left.\frac{\partial^2 S[\gamma+\lambda\xi+\mu\eta]}{\partial\lambda\partial\mu}\right|_{\substack{\lambda=0\\ \mu=0}} = \left[\frac{\partial}{\partial\lambda}\left(\frac{\partial}{\partial\mu}S[\gamma+\lambda\xi+\mu\eta]\right)\Big|_{\mu=0}\right]\Big|_{\lambda=0}$$
$$= \frac{\partial}{\partial\lambda}\int_a^b\left(\frac{\partial L}{\partial y^i}-\frac{d}{dt}\frac{\partial L}{\partial \dot y^i}\right)\eta^i dt\Big|_{\lambda=0},$$

其中 $y(t)=x(t)+\lambda\xi(t), L=L(y,\dot y)$. 像在定理 31.1 的证明同样的处理, 我们得到

$$\int_a^b\frac{\partial}{\partial\lambda}\left(\frac{\partial L}{\partial y^i}-\frac{d}{dt}\frac{\partial L}{\partial\dot y^i}\right)_{\lambda=0}\eta^i dt=\int_a^b\left(\frac{\partial M_i}{\partial y^k}\xi^k+\frac{\partial M_i}{\partial \dot y^k}\dot\xi^k+\frac{\partial M_i}{\partial\ddot y^k}\ddot\xi^k\right)_{\lambda=0}\eta^i dt,$$

其中

$$M_i=M_i(y,\dot y,\ddot y)=\frac{\partial L}{\partial y^i}-\frac{d}{dt}\frac{\partial L}{\partial\dot y^i}=\frac{\partial L}{\partial y^i}-\frac{\partial^2 L}{\partial y^j\partial\dot y^i}\dot y^j-\frac{\partial^2 L}{\partial\dot y^j\partial\dot y^i}\ddot y^j.$$

显式的计算导出下面的结果

$$\left.\frac{\partial^2 S[\gamma+\lambda\xi+\mu\eta]}{\partial\lambda\partial\mu}\right|_{\substack{\lambda=0\\ \mu=0}}=\int_a^b\left(\frac{\partial^2 L}{\partial x^i\partial x^j}\xi^j+\frac{\partial^2 L}{\partial x^i\partial\dot x^j}\dot\xi^j-\right.$$
$$\left.\frac{d}{dt}\left(\frac{\partial^2 L}{\partial\dot x^i\partial x^j}\xi^j+\frac{\partial^2 L}{\partial\dot x^i\partial\dot x^j}\dot\xi^j\right)\right)\eta^i dt.$$

引理得证. □

定义 36.1. 引理中作用于沿曲线 γ 的向量场 $\xi(t)$ 的线性算子 $J=(J_{ij})$ 被称为雅可比算子.

我们只给出欧拉—拉格朗日方程为测地方程时的例子. 此时适当选取的作用为

$$S=\int_a^b\frac{1}{2}g_{ij}\dot x^i\dot x^j dt \tag{6}$$

(并不是长度 $l=\int\sqrt{g_{ij}\dot x^i\dot x^j}dt$, 尽管它们有相同的极值线). 有下面定理.

定理 36.1. 对 $S=\int_a^b\frac{1}{2}g_{ij}\dot x^i\dot x^j dt$ 及任意测地线 $\gamma: x^i=x^i(t)$, 双线性形式 $G_\gamma(\xi,\eta)=\dfrac{\partial^2 S[\gamma+\lambda\xi+\mu\eta]}{\partial\lambda\partial\mu}$ 的形状为

$$G_\gamma(\xi,\eta)=-\int_a^b(\nabla^2_{\dot x}\xi^i+\dot x^j\dot x^k\xi^l R^i_{jkl})\eta^m g_{im}dt, \tag{7}$$

或者

$$G_\gamma(\xi,\eta)=-\int_a^b\langle J\xi,\eta\rangle dt, \tag{8}$$

其中

$$(J\xi)^i=\nabla^2_{\dot x}\xi^i+\dot x^j\dot x^k\xi^l R^i_{jkl}, \tag{9}$$

R^i_{jkl} 为曲率张量,t 为沿测地线 γ 的自然参数.

证明 对于拉格朗日 $L=\frac{1}{2}g_{ij}\dot{x}^i\dot{x}^j$ 的一阶变分为 (参见 §31)

$$\left.\frac{\partial}{\partial\lambda}S[\gamma+\lambda\xi]\right|_{\lambda=0}=-\int_a^b(\ddot{x}^k+\Gamma_{ij}^k\dot{x}^i\dot{x}^j)g_{kl}\xi^l dt$$
$$=-\int_a^b\langle\nabla_{\dot{x}}\dot{x},\xi\rangle dt.$$

因此,

$$\left.\frac{\partial^2}{\partial\lambda\partial\mu}S[\gamma+\lambda\xi+\mu\eta]\right|_{\substack{\lambda=0\\ \mu=0}}=-\frac{\partial}{\partial\lambda}\int_a^b\langle\nabla_{\dot{y}}\dot{y},\eta\rangle dt\Big|_{\lambda=0}$$
$$=-\int_a^b\{\langle\nabla_\xi\nabla_{\dot{x}}\dot{x},\eta\rangle dt+\langle\nabla_{\dot{x}}\dot{x},\nabla_\xi\eta\rangle\}dt$$

(这里的 $y(t)=x(t)+\lambda\xi(t)$). 由于测地线方程为 $\nabla_{\dot{x}}\dot{x}=0$, 故积分号内的第二项在曲线 γ 为零. 变换第一项. 我们有

$$\nabla_\xi\nabla_{\dot{x}}-\nabla_{\dot{x}}\nabla_\xi=R(\dot{x},\xi),\quad \nabla_\xi\dot{x}=\nabla_{\dot{x}}\xi$$

(因为联络的对称性). 最后得出

$$\nabla_\xi\nabla_{\dot{x}}\dot{x}=\nabla_{\dot{x}}^2\xi+R(\dot{x},\xi)\dot{x},$$

从而给出了定理. □

例 36.1. 考虑二维情形. 在测地线 $\gamma(t)$ 附近引进特别的坐标系 $(x,y),x=x^1,y=x^2$, 使得

a) 曲线 $(x,0)$ 自身即测地线,x 为自然参数;

b) 坐标 y 与测地线正交, 这时 $g_{ij}(x,0)=\delta_{ij}$.

这时对于 $\xi(t),\eta(t)$ 的双线性形式 $G_\gamma(\xi,\eta)$ 有形式

$$G_\gamma(\xi,\eta)=-\int_a^b\left(\frac{d^2}{dt^2}\xi^i+K(t)\xi^i\right)\eta_i dt,\tag{10}$$

其中 $\xi(t),\eta(t)$ 正交于 $\gamma(t)$, K 为高斯曲率.(为确信于此, 必须在 $G_\gamma(\xi,\eta)$ 的一般公式中代入关系式 $\nabla_{\dot{x}}^2(\xi)=\ddot{\xi}$(其正确性可由 $\Gamma_{j1}^i=-\Gamma_{i1}^j$ 在 γ 上成立得到) 并考虑到 $R_{1212}=K$(参看 §31.3).) 我们注意到, 如果向量场中至少有一个与切向量 $\dot{x}(t)$ 以常数因子成比例, 则会有 $G_\gamma(\xi,\eta)\equiv 0$.

注 二阶变分公式可以被推广到 "间断向量场", 这时导数 $\nabla_{\dot{x}}\xi$ 可以有间断点.

习题 36.1. 证明在间断向量场的情形下, 双线性形式 $G_\gamma(\xi,\eta)$ 具形式

$$G_\gamma(\xi,\eta)=-\sum_{P_i}\langle\eta,\Delta_{P_i}(\nabla_{\dot{x}}\xi)\rangle-\int_a^b\langle J\xi,\eta\rangle dt,\tag{11}$$

其中 $\Delta_P(\nabla_{\dot{x}}\xi)$ 为在点 P 这个协变导数的跃变, 而取和是对 $\nabla_{\dot{x}}\xi$ 的所有间断点进行的.

2. 共轭点和极小性条件

我们假定度量 g_{ij} 是黎曼的. 测地线的极小性条件由二次型 $G_\gamma(\xi,\xi)$ 在向量场上的正定性, 以及这些向量场在测地线 γ 的终点为零构成. 我们首先来解释何时双线性形式 $G_\gamma(\xi;\eta)$ 非异.

定义 36.2. 称沿从 P 到 Q 的极值曲线 γ 的向量场 ξ 是雅可比向量场是说, 它是雅可比方程的解:$J\xi=0$, 且在端点 P 和 Q 为零. 特别, 对于所研究的情形 $S=\dfrac{1}{2}\displaystyle\int\langle\dot{x},\dot{x}\rangle dt$, 我们有

$$J\xi=(\nabla_{\dot{x}}^2\xi^i+\dot{x}^j\dot{x}^k\xi^l R^i_{jkl})=0,\quad i=1,\cdots,n. \tag{12}$$

定义 36.3. 称点 P 和 Q 为沿从 P 到 Q 的测地线 γ 共轭的是说, 存在非零的沿曲线 γ 的雅可比向量场 ξ.

引理 36.2. 双线性形式 $G_\gamma(\xi,\eta)$ 为非异当且仅当测地线 γ 的端点 P 和 Q 不是沿 γ 共轭的.

证明 我们记得, 称双线性形式 $G_\gamma(\xi,\eta)$ 为非异是说, 如果不存在那样的向量场, 使得对任意的 η 有 $G_\gamma(\xi;\eta)=0$ (这里的向量场 ξ,η 在端点 P 和 Q 化为零). 如果向量场 ξ 是雅可比的, 则按照定理 36.1, 对任意 η 均有 $G_\gamma(\xi;\eta)=0$. 反之, 设形式 G_γ 在向量场 ξ 退化, 即 $G_\gamma(\xi,\eta)$ 对任意 η 为零. 考虑 $\eta=\alpha(t)J\xi$, 其中的函数 $\alpha(t)$ 非负并且在端点 (即当 $t=a$ 和 $t=b$) 为零. 于是由二阶变分公式得到

$$G_\gamma(\xi,\eta)=-\int_a^b\alpha(t)\langle J\xi,J\xi\rangle dt\equiv 0.$$

由此得到 $J\xi=0$(这是因为 $\alpha(t)$ 是任意的, 而度量是黎曼的). 因此端点为共轭. 引理得证. □

我们现在证明极小性的一个重要的必要条件.

定理 36.2. 如果从点 P 到点 Q 的测地线 γ 包含了其上的一对共轭点 P',Q', 则此测地线 γ 不是极小.

证明 (P,Q 沿 γ 不共轭的情形). 这时双线性形式 $G_\gamma(\xi,\eta)$ 非异. 故而对于极小测地线二次型 $G_\gamma(\xi,\xi)$ 应该为正定:如果 ξ 不恒为零, 则 $G_\gamma(\xi;\xi)>0$. 我们来证明, 当存在共轭点时这个要求不满足. 设 ξ' 为对应于点 P',Q' 的雅可比场.我们来构造 P 和 Q 之间的 "间断场" ξ: 在 P' 和 Q' 之间等于 ξ', 而在曲线段 $P'Q'$ 外为零 (图 37). 于是由公式 (11)

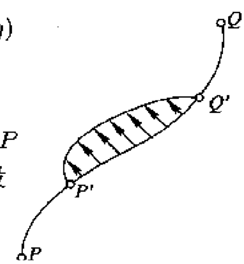


图 37

$$G_\gamma(\xi;\xi)=0.$$

这与形式 G_γ 的正定性矛盾. 定理得证. □

定理 36.3. 在充分小的测地线的长度范围内, 在所有连接相同两点的光滑曲线中, 作用泛函 $S[\gamma]$ 给出了极小. 因此每条充分短的测地线局部地实现了两点之间的距离.

证明 测地线 $\gamma(t)$ 的极小性条件为二次型 $G_\gamma(\xi,\xi)$ 在端点为 0 的所有向量场 ξ 上为正定. 由公式 (8)

$$\begin{aligned} G_\gamma(\xi,\xi) &= -\int_a^b [\langle\nabla_{\dot x}^2\xi,\xi\rangle + \langle R(\dot x,\xi)\dot x,\xi\rangle]dt \\ &= -\int_a^b \nabla_{\dot x}\langle\nabla_{\dot x}\xi,\xi\rangle dt + \int_a^b [\langle\nabla_{\dot x}\xi,\nabla_{\dot x}\xi\rangle - \langle R(\dot x,\xi)\dot x,\xi\rangle]dt \\ &= \int_a^b [\langle\nabla_{\dot x}\xi,\nabla_{\dot x}\xi\rangle - \langle R(\dot x,\xi)\dot x,\xi\rangle]dt \end{aligned}$$

(分部积分, 并考虑等式 $\xi(a)=\xi(b)=0$). 对于充分小的长度区间 Δl, 成立估值

$$\left|\int_a^b \langle R(\dot x,\xi)\dot x,\xi\rangle dt\right| < c(\Delta l)\int_a^b \langle\nabla_{\dot x}\xi,\nabla_{\dot x}\xi\rangle dt, \tag{13}$$

其中 $c(\Delta l)$ 为某个以度量 g_{ij} 和长度 Δl 表示的常数, 同时 $c(\Delta l)$ 当 $\Delta l\to 0$ 时趋向于零. 因为

$$\int_a^b \langle\nabla_{\dot x}\xi,\nabla_{\dot x}\xi\rangle dt > 0,$$

由此得到了我们的论断. □

习题

36.2. 证明不等式 (13).

(提示: 如果 $\xi(a)=\xi(b)=0$, 则 $|\xi| <$ 常数 $\cdot \max\limits_{[a,b]}|\nabla_{\dot x}\xi|\Delta t$, 其中 $\Delta l = |b-a|$.)

第六章　高维变分问题. 场及几何不变量

§37. 最简单的高维变分问题

1. 欧拉 – 拉格朗日方程

考虑在欧氏空间 $\mathbb{R}^n$ 中的区域 D, 其中 $\mathbb{R}^n$ 具有欧氏坐标 $x^1,\cdots,x^n$, 而区域 D 具有光滑 (或逐段光滑) 的边界 ∂D, 又考虑在 D 上的光滑向量值函数的线性空间 $F: F=\{f(x^1,\cdots,x^n)|f=(f^1,\cdots,f^k)\}$. 称区域 D 为变量 $x^1,\cdots,x^n$ 的定义域. 设 $L(x^i;p^\beta;q_j^\alpha)$ 为三组变量的光滑函数:$x^i, 1\leqslant i\leqslant n; p^\beta, 1\leqslant\beta\leqslant k; q_j^\alpha, 1\leqslant\alpha\leqslant k; 1\leqslant j\leqslant n$. 称函数 L 为拉格朗日, 并构造定义在 F 上的泛函 $I[f]$

$$I[f]=\int\limits_D L(x^i;f^\beta(x^i);f_{x^j}^\alpha(x^i))dx^1\wedge\cdots\wedge dx^n,$$

其中 $\int\limits_D$ 表示在 n 维区域 D 上的重积分 $\int\cdots\int\limits_D$ (为简单起见, 我们假定区域 D 在 $\mathbb{R}^n$ 中有界) , $dx^1\wedge\cdots\wedge dx^n=d^nx$ 为 $\mathbb{R}^n$ 中的 n 维体积形式; $f_{x^j}^\alpha(x^i)=\dfrac{\partial}{\partial x^j}(f^\alpha(x^i))$. 简短地我们记 $I[f]=\int\limits_D L(x^i;f^\beta;f_{x^j}^\alpha)d^nx$.

最简单的情形:$n=1$ (一维变分问题) 已在前面考虑过了: 例如弧长函数 $l(\gamma)=\int_0^l\sqrt{g_{\alpha\beta}(y)\dot{y}^\alpha\dot{y}^\beta}dt$, 路径 $\gamma(t)$ 上的作用泛函 $S(\gamma)=\int_0^l g_{\alpha\beta}(y)\dot{y}^\alpha\dot{y}^\beta dt$, 它定义在黎曼空间中的分段光滑曲线 $\gamma(t)=(y^1(t),\cdots,y^k(t))$ 上.

在这一章里我们的兴趣在高维的泛函:$n>1$. 最简单的例子是在二维曲面上的

面积泛函 $I[f]=\iint\limits_D \sqrt{EG-F^2}dxdy$, 其中 $D\subset\mathbb{R}^2(x,y)$ 为参数 x,y 变动的区域; $f(x,y)=(u^1(x,y),u^2(x,y),u^3(x,y))$ 为在 $\mathbb{R}^3(u^1,u^2,u^3)$ 中的曲面, 其诱导度量为 $dl^2=Edx^2+2Fdxdy+Gdy^2$(参看 §7.3).

$$L(x,y;f;f_x,f_y)=L(f_x,f_y)=\sqrt{EG-F^2}=\sqrt{\langle f_x,f_x\rangle\langle f_y,f_y\rangle-\langle f_x,f_y\rangle^2}.$$

我们转向考虑一般形式的泛函.

考虑 "点"$f\in F$ (F 为泛函 I 的定义域). 设 η 为在区域 D 上的光滑有限函数, 即在 D 中的开集 (其闭包在 D 中) 的外部为零的函数, 特别地, 可假定在区域 D 的边界 ∂D 上 $\eta=0$. 一般地就我们的目的而言, 只要考虑具充分小支集的函数就够了. 如果对函数 f, 有 $f+\varepsilon\eta\in F$, 我们则称这个函数 η 为函数 f 的扰动. 这里的 ε 为小参数. 现在构造表达式 $\frac{1}{\varepsilon}(I[f+\varepsilon\eta]-I[f])$ 并取极限 (按 ε), 便得到函数

$$\left.\frac{dI[f+\varepsilon\eta]}{d\varepsilon}\right|_{\varepsilon=0}=\int\frac{\delta I}{\delta f}\delta f d^nx,\quad \delta f=\eta,\tag{1}$$

自然地会把它叫做 "泛函 I 在点 f 按方向 η 的导数".(1) 中定义的向量 $\frac{\delta I}{\delta f}=\left(\frac{\delta I}{\delta f^\beta}\right)$ 被称做泛函 $I[f]$ 的**变分导数**.

定义 37.1. 称函数 $f_0\in F$ 是对泛函 I 的定常(或者极值, 或者判别) 函数是说有 $\frac{\delta I[f_0]}{\delta f}\equiv 0$, 即对任意 $\eta=\delta f$ 使 $f+\varepsilon\eta\in F$ 的有限扰动表达式 (1) 为 0.

我们对导数 $\frac{\delta I[f_0]}{\delta f}$ 进行分析和研究.

$$\Delta I[f]=\int\limits_D[L(x^i;f^\beta+\varepsilon\eta^\beta;f^\alpha_{x^j}+\varepsilon\eta^\alpha_{x^j})-L(x^i;f^\beta;f^\alpha_{x^j})]d^nx.$$

展开被积表达式为泰勒级数, 我们得到

$$\begin{aligned}\Delta I[f]&=\int\limits_D\left[\sum_{\beta=1}^k\frac{\partial L}{\partial f^\beta}\varepsilon\eta^\beta+\sum_{\alpha=1}^k\sum_{j=1}^n\frac{\partial L}{\partial f^\alpha_{x^j}}\varepsilon\eta^\alpha_{x^j}+o(\varepsilon)\right]d^nx\\&=\varepsilon\int\limits_D\left[\sum_{\alpha=1}^k\left[\frac{\partial L}{\partial f^\alpha}\eta^\alpha+\sum_{j=1}^n\frac{\partial L}{\partial f^\alpha_{x^j}}\eta^\alpha_{x^j}\right]d^nx+\int\limits_D o(\varepsilon)d^nx.\right.\end{aligned}$$

进行分部积分. 我们得到

$$\Delta I[f] = \varepsilon \int\limits_D \sum_{j=1}^{n} \frac{\partial}{\partial x^j}\left(\frac{\partial L}{\partial f^{\alpha}_{x^j}}\eta^{\alpha}\right) d^n x + \varepsilon \int\limits_D \sum_{\alpha=1}^{k}\left(\frac{\partial L}{\partial f^{\alpha}} - \sum_{j=1}^{n}\frac{\partial}{\partial x^j}\left(\frac{\partial L}{\partial f^{\alpha}_{x^j}}\right)\right)\eta^{\alpha} d^n x + \int\limits_D o(\varepsilon) d^n x.$$

在第一个积分中, 对变量 x^j 的积分可以从对其余变量 $x^1, \cdots, \widehat{x}^j, \cdots, x^n$ 的积分中分离出来. 我们有

$$\int\limits_D \frac{\partial}{\partial x^j}\left(\frac{\partial L}{\partial f^{\alpha}_{x^j}}\eta^{\alpha}\right) d^n x = \int\limits_{x^1,\cdots,\widehat{x}^j,\cdots,x^n}\left[\int_P^Q \frac{\partial}{\partial x^j}\left(\frac{\partial L}{\partial f^{\alpha}_{x^j}}\eta^{\alpha}\right) dx^j\right] d^{n-1}x$$

(P 和 Q 依赖于 $x^1, \cdots, \widehat{x}^j, \cdots, x^n; d^{n-1}x = dx^1 \wedge \cdots \wedge d\widehat{x}^j \wedge \cdots \wedge dx^n$). 在内层的积分 $\int_P^Q$ 对变量 $x^1, \cdots, \widehat{x}^j, \cdots, x^n$ 可以看作是对参数的积分, 故有

$$\int\limits_D \frac{\partial}{\partial x^j}\left(\frac{\partial L}{\partial f^{\alpha}_{x^j}}\eta^{\alpha}\right) d^n x = \int\limits_{x^1,\cdots,\widehat{x}^j,\cdots,x^n}\left[\frac{\partial L}{\partial f^{\alpha}_{x^j}}\eta^{\alpha}\bigg|_Q - \frac{\partial L}{\partial f^{\alpha}_{x^j}}\eta^{\alpha}\bigg|_P\right] d^{n-1}x \equiv 0,$$

这是因 $\eta^{\alpha}(Q) = \eta^{\alpha}(P) = 0$ (参看前面). 因此

$$\Delta I[f] = \varepsilon \int\limits_D \sum_{\alpha=1}^{k}\left[\frac{\partial L}{\partial f^{\alpha}} - \sum_{j=1}^{n}\frac{\partial}{\partial x^j}\left(\frac{\partial L}{\partial f^{\alpha}_{x^j}}\right)\right)\eta^{\alpha} d^n x + \int\limits_D o(\varepsilon) d^n x.$$

因此, 因为 $\lim\limits_{\varepsilon \to 0}\int\limits_D o(\varepsilon) d^n x = 0$, 便有

$$\frac{\delta I[f]}{\delta f^{\alpha}} = \frac{\partial L}{\partial f^{\alpha}} - \sum_{j=1}^{n}\frac{\partial}{\partial x^j}\left(\frac{\partial L}{\partial f^{\alpha}_{x^j}}\right). \tag{2}$$

如果 $f_0 \in F$ 为对 $I[f]$ 的定常 (极值) 函数, 则对任意有限扰动 η 满足恒等式

$$\int\limits_D \sum_{\alpha=1}^{k}\eta^{\alpha}\left[\frac{\partial L}{\partial f^{\alpha}} - \sum_{j=1}^{n}\frac{\partial}{\partial x^j}\left(\frac{\partial L}{\partial f^{\alpha}_{x^j}}\right)\right]\bigg|_{f=f_0} d^n x = 0,$$

于是

$$\frac{\delta I[f]}{\delta f^{\alpha}} = \frac{\partial L}{\partial f^{\alpha}_0} - \sum_{j=1}^{n}\frac{\partial}{\partial x^j}\left(\frac{\partial L}{\partial f^{\alpha}_{0,x^j}}\right) = 0 \ (1 \leqslant \alpha \leqslant k). \tag{3}$$

称微分方程组 (3) 为对泛函 I 的欧拉 – 拉格朗日方程组.

把前面的结果写成定理的形式.

定理 37.1. 函数 $f_0 \in F$ 对泛函 $I[f]$ 为定常的充要条件是它满足欧拉 – 拉格朗日方程组 (3).

2. 能量 – 动量张量

我们考虑形如 $I[f]=\int_D L(f^\beta, f^\alpha_{x^i})d^n x$ 的泛函, 其中的拉格朗日函数 L 不显式地含有变量 $x^i, 1\leqslant i\leqslant n$. 极值函数 f (被说成是该"系统的运动") 为欧拉 – 拉格朗日方程组的解, 此方程组为

$$\frac{\partial L}{\partial f^\alpha}-\frac{\partial}{\partial x^j}\left(\frac{\partial L}{\partial f^\alpha_{x^j}}\right)=0. \tag{4}$$

像在力学中按照能量守恒定律处理问题那样我们来进行类似的处理, 我们来处理下面的变换:

$$\frac{\partial L}{\partial x^i}=\frac{\partial L}{\partial f^\alpha}\frac{\partial f^\alpha}{\partial x^i}+\frac{\partial L}{\partial f^\alpha_{x^j}}\frac{\partial(f^\alpha_{x^j})}{\partial x^i}.$$

在这里代入 (4) 得到

$$\frac{\partial L}{\partial x^i}=\frac{\partial}{\partial x^j}\left(\frac{\partial L}{\partial f^\alpha_{x^j}}\right)\frac{\partial f^\alpha}{\partial x^i}+\frac{\partial L}{\partial f^\alpha_{x^j}}\frac{\partial(f^i_{x^\alpha})}{\partial x^\beta}.$$

因为 $\dfrac{\partial}{\partial x^i}(f^\alpha_{x^j})=\dfrac{\partial^2 f^\alpha}{\partial x^i\partial x^j}=\dfrac{\partial}{\partial x^j}(f^\alpha_{x^i})$, 则

$$\frac{\partial L}{\partial x^i}=\frac{\partial}{\partial x^j}\left(\frac{\partial L}{\partial f^\alpha_{x^j}}\right)\frac{\partial f^\alpha}{\partial x^i}+\frac{\partial L}{\partial f^\alpha_{x^j}}\frac{\partial}{\partial x^j}(f^\alpha_{x^i})=\frac{\partial}{\partial x^j}\left(\frac{\partial L}{\partial f^\alpha_{x^j}}f^\alpha_{x^i}\right).$$

同时 $\dfrac{\partial L}{\partial x^i}=\delta^j_i\dfrac{\partial L}{\partial x^j}$, 因此

$$\frac{\partial L}{\partial x^j}\delta^j_i=\frac{\partial}{\partial x^j}\left(\frac{\partial L}{\partial f^\alpha_{x^j}}f^\alpha_{x^i}\right),$$

即

$$\frac{\partial}{\partial x^j}\left(\delta^j_i L-f^\alpha_{x^i}\frac{\partial L}{\partial f^\alpha_{x^j}}\right)=0. \tag{5}$$

引入记号

$$T^j_i=f^\alpha_{x^i}\frac{\partial L}{\partial f^\alpha_{x^j}}-\delta^j_i L. \tag{6}$$

于是关系 (5) 有形式

$$\frac{\partial T^j_i}{\partial x^j}=0. \tag{7}$$

这个方程意味着在整个区域 D 上, 张量 (T^j_i) 的发散量等于零.

注 在欧氏坐标下张量的上与下指标没有区别. 而在伪欧氏空间中, 如其度量为 g_{ij}, 则可以定义张量 T_{ij} 为

$$T_{ij}=g_{jl}f^\alpha_{x^i}\frac{\partial L}{\partial f^\alpha_{x^l}}-g_{ij}L=g_{il}T^l_i. \tag{8}$$

定义 37.2. 称张量 T^j_i 为所给系统的能量 – 动量张量(具拉格朗日 $L(f,f_{x^j})$).

方程 $\dfrac{\partial T_i^j}{\partial x^j}=0$ 并不唯一确定能量 - 动量张量. 事实上, 由公式 (6) 或 (8) 定义的张量 T_i^j 可以附加上形如 $\dfrac{\partial}{\partial x^l}\psi_i^{jl}$ 的张量和, 其中 ψ_i^{jl} 为任意对指标 j,l 为反称的张量. 于是显然地, 新张量 $\tilde{T}_i^j=T_i^j+\dfrac{\partial\psi_i^{jl}}{\partial x^l}$ 也满足关系式 (7), 其原因是 $\dfrac{\partial^2\psi_i^{jl}}{\partial x^j\partial x^l}=0$.

一般说来, 上面所定义的张量 T_{ij} 不是对称的, 但是如果以方程 (7) 的观点考虑它, 则可以借助于加上相应反称 (对 i,j) 的形如 $\sum\limits_i\dfrac{\partial}{\partial x^i}(\psi_{ij}^l)$ 张量可以把它做成为对称的 (并保持 (7) 不变). (只需要求使张量 ψ_{ij}^l 对 i,j 为反称就可以了, 即要使张量 ψ_{ijl} 对所有指标为反称.) 作为对称张量的能量 - 动量张量, 其唯一性的定义在某些物理问题中起了特别的作用 (参看下面).

在后面的能量 - 动量张量概念的基本应用集中在四维伪欧氏空间中的变分问题, 特别是闵可夫斯基空间 $\mathbb{R}_1^4$, 其坐标为 x^0,x^1,x^2,x^3, 而 x^0 与 t (时间) 成比例, 而 x^1,x^2,x^3 为空间的坐标.

引进具拉格朗日 L 的动量系统的 4 维向量. 为此我们要用到能量 - 动量张量 T_{ij}.

设 $dS_k=\dfrac{1}{6}\varepsilon_{ijmk}dx^i\wedge dx^j\wedge dx^m$ 是在 3 维超平面坐标下的标准形式 (参看习题 26.2).

定义 37.3. 称向量 $P=(P_0,P_1,P_2,P_3)$ 为具拉格朗日 L 的动量 4 元向量 (具下指标)是说其中

$$P_i=\lambda\int\limits_{x^0=\text{常数}}T_i^kdS_k=\lambda\int\limits_{x^0=\text{常数}}T_i^0dS_0,\quad i=0,1,2,3,\lambda\text{为常数}.$$

(在这里我们用到了: 如 $k\neq 0$, 则形式 dS_k 在超平面 $x_0=$ 常值的限制等于零.)

注 按照类似于 (31.7) 的能量公式, 分量 $T_0^0=\dot{f}^j\dfrac{\partial L}{\partial\dot{f}^j}-L$ 应被看成是能量密度(这里的 "·" 表示对 $x^0=ct$ 的导数). 因此 $\int T_0^0d^3x$ ($d^3x=dx^1\wedge dx^2\wedge dx^3$) 为系统的总能量. 然而我们由定义 37.3, 对于时间分量 P_0 有下面的表达式: $P_0=\lambda\int\limits_{x^0=\text{常数}}T_0^0d^3x$. 通常令 $\lambda=\dfrac{1}{c}$, 于是分量 P_0 与系统的总能量相同, 只差一个因子 $1/c$ (如同相对论的质点的动量 4 元向量 —— 见 §32).

命题 37.1. 条件 $\dfrac{\partial T_i^k}{\partial x^k}=0$ 等价于动量向量 P 的守恒条件.

证明 我们将假定, 所考虑的欧拉 - 拉格朗日方程的解, 即向量场 f, 在一个足够大的三维空间 (x^1,x^2,x^3) 中的球外为零 (或者比 $1/r^2$ 还要快地趋向于 0, 其中

$r^2=(x^1)^2+(x^2)^2+(x^3)^2$. 我们考虑在闵可夫斯基空间中的柱面 C, 其底面为 D_1 和 D_2, 分别由方程 $x^0=x_1^0$ 和 $x^0=x_2^0$ 给出, 而半径 R 足够大. 设 Π 为柱的侧曲面. 由条件 $\dfrac{\partial T_i^k}{\partial x^k}=0$ 和斯托克斯公式得到对每个 i 都有 $\displaystyle\int_C T_i^k dS_k=0$ (参看习题 26.2). 故而

$$\Big(\int\limits_{D_2}-\int\limits_{D_1}+\int\limits_{\Pi}\Big)T_i^k dS_k=0.$$

在侧曲面上的积分可以不予考虑 (当柱的半径 R 趋向无穷大时它趋向零). 当 $R\to\infty$ 我们得到

$$P_i(x_1^0)=\int\limits_{x^0=x_1^0}T_i^k dS_k=\int\limits_{x^0=x_2^0}T_i^k dS_k=P_i(x_2^0).$$

命题得证. □

引理 37.1. 系统的动量向量在 T_i^k 到 $\widetilde{T}_i^k=T_i^k+\sum\limits_l\dfrac{\partial\psi_i^{kl}}{\partial x^l}$ 的变换下不变, 其中的 ψ_i^{kl} 为对 k,l 反称的向量.

证明 分量 P_i 有形式 $\lambda\displaystyle\int\limits_{x^0=\text{常数}}T_i^k dS_k$, 因此, 需要证明 $\displaystyle\int\frac{\partial\psi_i^{kl}}{\partial x^l}dS_k=0$. 由于 ψ_i^{kl} 对于 k,l 的反对称性, 我们有

$$\int\limits_S\frac{\partial\psi_i^{kl}}{\partial x^l}dS_k=\frac{1}{2}\int\limits_S\left(dS_k\frac{\partial\psi_i^{kl}}{\partial x^l}-dS_l\frac{\partial\psi_i^{kl}}{\partial x^k}\right).$$

右端的积分形状是 $\displaystyle\int\limits_S d\omega_i^{(2)}$, 其中 $\omega_i^{(2)}$ 为在三维曲面 $S:x^0=$ 常数上的一个二次形式. 由斯托克斯公式 $\displaystyle\int\limits_S d\omega_i^{(2)}=\int\limits_{\partial S}\omega_i^{(2)}$, 其中 ∂S 为在 S 中的二维闭曲面 ("离开无穷远处" 的边界), 然后, 假定分量 ψ_i^{kl} 以充分快的速度 "在无穷远处" 趋向于零, 那么我们便得到 $\displaystyle\int\limits_{\partial S}\omega_i^{(2)}=0$. 引理得证. □

如果我们有一个系统, 其拉格朗日为 L, 那么它的能量 – 动量张量 $T^{ik}=g^{ij}T_j^k$ 在所感兴趣的物理例子中, 可以被做成对称的张量而不改变动量. 现在设张量 T^{ik} 为对称.

由张量 T^{ik} 的对称性, 我们推导出被称做的 "动量矩守恒定律".

定义 37.4. 称张量

$$M^{ik}=\int(x^i dP^k-x^k dP^i)=\frac{1}{c}\int\limits_{x^0=\text{常数}}(x^iT^{kl}-x^kT^{il})dS_l$$

为动量矩.

这个积分是在三维区域 S 上进行的. 这个公式是对 §32 中质点组所计算的经典动量矩公式

$$M^{ik}=\sum(P^kx^i-P^ix^k)$$

的自然推广, 其中取和是对系统中的所有质点的.

引理 37.2. 如果能量 - 动量张量 T^{ik} 为对称, 则动量矩的张量 M^{ik} 守恒①.

证明 像上面那样, 我们应该证明

$$M^{ik}(x_1^0)=\frac{1}{c}\int\limits_{D_1}(x^iT^{kl}-x^kT^{il})dS_l$$

与

$$M^{ik}(x_2^0)=\frac{1}{c}\int\limits_{D_2}(x^iT^{kl}-x^kT^{il})dS_l$$

相等, 其中 D_1 和 D_2 例如由方程 $x^0=x_1^0, x^0=x_2^0$ 给出. 由于前面所做的注解及斯托克斯公式就足以让我们相信, 在 M^{ik} 中被积表达式的发散量应该等于零

$$\frac{\partial}{\partial x^l}(x^iT^{kl}-x^kT^{il})=0.$$

计算此导数, 因为张量 T^{ik} 的对称性和方程 $\dfrac{\partial T^{kl}}{\partial x^l}=0$, 我们有

$$\frac{\partial}{\partial x^l}(x^iT^{kl}-x^kT^{il})=T^{ki}-T^{ik}=0.$$

引理得证. □

转向高维变分问题的具体例子.

3. 电磁场方程

电磁场方程 (麦克斯韦方程) 就是对于泛函作用 $S=S_f+S_m+S_{mf}$ 的欧拉 - 拉格朗日方程, 我们现在把它们描述出来. S_m 这一项是那样的部分作用, 它只由那些在电磁场中运动的质点 (电荷) 所决定 (即 S_m 是在没有场的情形下质点的作用). 通常这个作用是这样确定的: $S_m=-\sum mc\int_a^b dl$, 其中的和取遍所有在场中运动的

①更准确地说, 如果能量 - 动量张量对称, 则普通物理学定义的动量矩守恒可化为上面提及的直观公式.

质点,m 为质点的质量,c 为光速, 积分 $\int_a^b dl$ 则沿着两个固定事件之间的在 $\mathbb{R}^4_{1,3}$ 中的世界线: 质点在时刻 t_1, t_2 所处的初始位置和终止位置; l 为弧长. 这个作用可以表示成三维形式: $S_m = \int_{t_1}^{t_2} L dt$, 其中 L 为拉格朗日函数: $L = -mc^2\sqrt{1-v^2/c^2}$, v 为三维速度 (参看 §32).

S_{mf} 这一项是那种部分作用, 它由场中质点相互作用产生的. 通常这作用是这样决定的: $S_{mf} = -\sum\limits_{(j)} \frac{e_j}{c}\int A_k^{(j)} dx^k$, 这里的和取遍有具电荷 e_j 的质点 j, 积分则跑过质点的世界线 (就像前面的那些积分一样). 4 元向量 (A_i) 刻画了这个场的特征, 通常称它为 4 元势能; 它的分量 A_i 依赖于时间和空间坐标. 在 S_{mf} 的积分中, 向量 (A_i) 的值是在质点 j 的世界线上被赋值的, 并沿此曲线进行积分. 以它与电磁场相互作用的观点看, 质点 j 的性质只被一个参数决定, 即电荷 e_j, 它包含于作用 S_{mf} 中. 因此对于电磁场中的电荷, 作用 $S_m + S_{mf}$ 的形式为

$$\int_a^b \left(-mcdl - \frac{e}{c}A_k dx^k\right).$$

S_f 这一项是那样的部分作用, 它只依赖于场自身的性质; 换句话说,S_f 是在无电荷时场的作用. 如果我们只对电荷在所给电磁场中的运动有兴趣, 那么项 S_f 则不是实质性的, 但是只要我们提出求由此场决定的方程这些问题时, 这一项则是本质的了.

我们记得,4 元向量 (A^i) 的三个空间分量 (A^1, A^2, A^3) 构成了一个三维向量 A, 它称做场的*向量势能*. 时间分量 A^0 有形式 $A^0 = \varphi$, 这里的实函数 φ 称做场的*标量势能*.这里的指标可用闵可夫斯基度量将其下降.

我们还记得, 向量 $\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} - \operatorname{grad}\varphi$ 叫做电场强度 (这是个三维向量), 另外称三维向量 $\mathbf{H} = \operatorname{rot}\mathbf{A}$ (按定义) 为磁场强度. 称 $\mathbf{E} \neq 0, \mathbf{H} = 0$ 的电磁场为电场, 而称 $\mathbf{E} = 0, \mathbf{H} \neq 0$ 时为*磁场*.

像通常那样, 以 $F_{ik} = \frac{\partial A_k}{\partial x^i} - \frac{\partial A_i}{\partial x^k}$ 表示电磁场张量. 于是作用 S_f(见前面) 的样子是

$$S_f = a\int 2(E^2 - H^2)d^4x,$$

其中 $H^2 = \langle \mathbf{H}, \mathbf{H}\rangle, E^2 = \langle \mathbf{E}, \mathbf{E}\rangle$ 为 3 维向量的平方标量,a 为某个常数, 积分遍历整个空间坐标, 从而在整个三维空间上, 而变量 x^0 (与时间成比例) 取在两个固定的时刻之间. 我们记得 $F_{ik}^2 = F_{ik}F^{ik} = 2(H^2 - E^2)$ (F_{ik} 为反称张量). 常数 a 通常取为

$-\dfrac{1}{16c\pi}$. 于是

$$S_f = -\frac{1}{16c\pi}\int 2(H^2 - E^2)d^4x = -\frac{1}{16c\pi}\int F_{ik}^2 d^4x.$$

因此, 对具有电荷的电场的整个作用 S 为

$$S = -\sum\int mcdl - \sum\int\frac{e}{c}A_k dx^k - \frac{1}{16c\pi}\int F_{ik}^2 d^4x. \tag{9}$$

到现在为止, 我们把电荷看作是点状的, 但常常方便地将电荷假设为在整个空间上为连续分布的状态. 如果 ρ 为电荷密度, 于是 ρdV 为在体积 $dV = d^3x$ 上所具有的电荷 (在三维空间中), ρ 与 x^1, x^2, x^3 和时间 t 相关.

考虑电荷在 $\mathbb{R}_1^4$ 中的世界线, 它被时间 t 所参数化:$x^0 = ct, x^\alpha = x^\alpha(t), \alpha = 1,2,3$, 且设 $\left(\dfrac{dx^i}{dt}\right)$ 为电荷的速度 4 元向量. 于是称 4 元向量 (j^i) 为电流 4 元向量, 其中 $j^i = \rho\dfrac{dx^i}{dt}$.

这个 4 元向量的三个空间的坐标 j^1, j^2, j^3 构成了通常三维电流向量 $\mathbf{j} = \rho\mathbf{v}$, 其中 $\mathbf{v}$ 为所给点上的电荷速度. 分量 j^0 等于 $c\rho$. 通过直接计算可证明, 在三维电流向量 j^i 中的作用可表示为

$$S = -\sum\int mcdl - \frac{1}{c^2}\int A_i j^i d^4x - \frac{1}{16c\pi}\int F_{ik}^2 d^4x$$

(请验证!).

现在转向求场的方程. 这时我们假定已经知道了电荷的运动. 这表明在推导欧拉方程时, 变分只需在位势场中进行 (即位势 4 元向量 A_i) 就可以了.

因此, 因为电荷的轨线无变差, 即 $\delta S_m = 0$, 但在 S_{mf} 这一项中也不应对电流 j^i 进行变分, 这是因为 $j^i = \rho\dfrac{dx^i}{dt}$. 我们从而有

$$\begin{aligned}\delta S &= \delta(S_{mf} + S_f) = -\int\left(\frac{1}{c^2}j^i\delta A_i + \frac{1}{16c\pi}\delta(F_{ik}^2)d^4x\right.\\ &= -\int\frac{1}{c}\left[\frac{1}{c}j^i\delta A_i + \frac{1}{8\pi}F^{ik}\delta F_{ik}\right]d^4x = 0.\end{aligned}$$

因为 $F_{ik} = \dfrac{\partial A_k}{\partial x^i} - \dfrac{\partial A_i}{\partial x^k}$, 于是

$$\begin{aligned}\delta S &= -\int\frac{1}{c}\left\{\frac{1}{c}j^i\delta A_i + \frac{1}{8\pi}F^{ik}\delta\left(\frac{\partial A_k}{\partial x^i} - \frac{\partial A_i}{\partial x^k}\right)\right\}d^4x\\ &= -\int\frac{1}{c}\left\{\frac{1}{c}j^i\delta A_i + \frac{1}{8\pi}F^{ik}\frac{\partial}{\partial x^i}(\delta A_k) - \frac{1}{8\pi}F^{ik}\frac{\partial}{\partial x^k}(\delta A_i)\right\}d^4x.\end{aligned}$$

由于 $F_{ik}=-F_{ki}$, 则

$$\delta S=-\int\frac{1}{c}\left\{\frac{1}{c}j^i\delta A_i-\frac{1}{4\pi}F^{ik}\frac{\partial}{\partial x^k}(\delta A_i)\right\}d^4x.$$

用分部积分对第二项积分, 则变为

$$\delta S=-\int\frac{1}{c}\left\{\frac{1}{c}j^i\delta A_i+\frac{1}{4\pi}\delta A_i\frac{\partial F^{ik}}{\partial x^k}\right\}d^4x+\frac{1}{4\pi c}\int F^{ik}\delta A_i dS_k\Big|_P^Q.$$

在后面的项中积分是在四维区域的边界上进行的, 这时在空间坐标下此区域的边界是“无限的”, 但在其上场为零 (故而 $F^{ik}\equiv 0$); 时间的边界为两个固定的时刻 t_1, t_2; 在这些点上势能的变分 $\{\delta A_i\}$ 为零.

于是, 最后面的这一项为零. 所以 $\int\left(\frac{1}{c}j^i+\frac{1}{4\pi}\frac{\partial F^{ik}}{\partial x^k}\right)\delta A_i d^4x=0$. 由于变差 δA_i 的任意性, 故有

$$\frac{\partial F^{ik}}{\partial x^k}=-\frac{4\pi}{c}j^i. \tag{10}$$

这是欧拉 - 拉格朗日方程, 它是由作用 S 对位势场的变分所得到的. 我们用三维形式来写出所得到的四个方程 ($i=0,1,2,3$). 回忆张量 F^{ik} 的显式表示:

$$(F^{ik})=\begin{pmatrix}0 & -E_x & -E_y & -E_z\\ E_x & 0 & -H_z & H_y\\ E_y & H_z & 0 & -H_x\\ E_z & -H_y & H_x & 0\end{pmatrix}$$

(在坐标 $x^0=ct, x^1=x, x^2=y, x^3=z$ 下写出的张量, 参看 §21). 那么, 第一个方程 (当 $i=1$) 有形式

$$\frac{1}{c}\frac{\partial F^{10}}{\partial t}+\frac{\partial F^{11}}{\partial x}+\frac{\partial F^{12}}{\partial y}+\frac{\partial F^{13}}{\partial z}=-\frac{4\pi}{c}j^1;$$

由此 $-\frac{\partial H_z}{\partial y}+\frac{\partial H_y}{\partial z}+\frac{1}{c}\frac{\partial E_x}{\partial t}=-\frac{4\pi}{c}j_x$. 对 $i=2,3$ 的方程有类似的变化, 最后给出了

$$\operatorname{rot}\mathbf{H}=\frac{1}{c}\frac{\partial\mathbf{E}}{\partial t}+\frac{4\pi}{c}\mathbf{j}. \tag{11}$$

第零个方程 (当 $i=0$) 给出

$$\frac{\partial E_x}{\partial x}+\frac{\partial E_y}{\partial y}+\frac{\partial E_z}{\partial z}=\frac{4\pi}{c}c\rho,$$

即

$$\operatorname{div}\mathbf{E}=4\pi\rho. \tag{11'}$$

方程 (11),(11′) 构成了第二对麦克斯韦方程. 回忆第一对麦克斯韦方程是

$$\text{rot } \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t}, \tag{12}$$

$$\text{div } \mathbf{H} = 0. \tag{12'}$$

方程 (11)—(12) 定义了电磁场, 从而是基础电磁场方程.

我们现在在没有电荷的条件下, 求出电磁场的能量 - 动量张量的显式表达. 作用 S_f 的形式为 $-\dfrac{1}{16c\pi}\displaystyle\int F_{ik}^2 d^4x$, 故而以第 2 小节的术语表示时, 在这一节中的函数 L 为

$$L = -\frac{1}{16c\pi}F_{ik}^2 = -\frac{1}{16c\pi}\left(\frac{\partial A_l}{\partial x^k} - \frac{\partial A_k}{\partial x_l}\right)^2.$$

在这里势能分量 A_i 起了第 2 小节中量 f^β 的作用. 于是由能量 - 动量张量 (T_i^k) 的定义, 我们得到了

$$T_i^k = \frac{\partial A_l}{\partial x^i}\frac{\partial L}{\partial\left(\dfrac{\partial A_l}{\partial x^k}\right)} - \delta_i^k L.$$

我们来求 $\dfrac{\partial L}{\partial\left(\dfrac{\partial A_l}{\partial x^k}\right)}$. 为此来计算微分 dL. 我们有

$$\begin{aligned} dL &= -\frac{1}{8\pi c}F^{kl}\left(d\frac{\partial A_l}{\partial x^k} - d\frac{\partial A_k}{\partial x^l}\right) \\ &= -\frac{1}{8\pi c}\left(F^{kl}d\frac{\partial A_l}{\partial x^k} - F^{kl}d\frac{\partial A_k}{\partial x^l}\right) \\ &= -\frac{1}{4\pi c}F^{kl}d\frac{\partial A_l}{\partial x^k}. \end{aligned}$$

于是 $\dfrac{\partial L}{\partial\left(\dfrac{\partial A_l}{\partial x^k}\right)} = -\dfrac{1}{4\pi c}F^{kl}$, 因此

$$T_i^k = -\frac{1}{4\pi c}\frac{\partial A_l}{\partial x^i}F^{kl} + \frac{1}{16\pi c}\delta_i^k F_{lm}F^{lm}.$$

用闵可夫斯基度量 g^{ik} 提升指标 i, 我们得到

$$T^{ik} = -\frac{g^{im}}{4\pi c}\frac{\partial A_l}{\partial x^m}F^{kl} + \frac{1}{16\pi c}g^{ik}F_{lm}F^{lm}.$$

所得到的这个能量 - 动量张量目前还不对称. 在前面我们需提出过一个办法来消去张量 T^{ik} 的反称部分. 把这个办法用到现在的情形时, 我们在 T^{ik} 中加入形如 $\dfrac{1}{4\pi c}\dfrac{\partial A^i}{\partial x^l}F^{kl}$ 的和, 它可以被表示为 $\dfrac{\partial}{\partial x^l}(\psi^{ikl})$ (见前面). 事实上, 因为在没有电荷

$(j=(j^i)=0)$ 时的麦克斯韦方程, 有 $\dfrac{\partial F^{kl}}{\partial x^l}=0$, 故

$$\frac{\partial A^i}{\partial x^l}F^{kl}=\frac{\partial}{\partial x^l}(A^iF^{kl})-A^i\frac{\partial F^{kl}}{\partial x^l}=\frac{\partial}{\partial x^l}(A^iF^{kl}).$$

(像在第 2 小节所证明的, 量 $\dfrac{\partial}{\partial x^l}(\psi^{ikl})$ 可以加到能量 – 动量张量上而不改变系统的动量向量.)

因为 $\dfrac{\partial A_l}{\partial x^i}-\dfrac{\partial A_i}{\partial x^l}=F_{il}$, 故最后我们得到了电磁场的对称的能量 – 动量张量:

$$T^{ik}=\frac{1}{4\pi c}(-F^{il}F_l^k+\frac{1}{4}g^{ik}F_{lm}F^{lm}). \tag{13}$$

习题 37.1. 设电荷密度 ρ 等于零, 电磁场 A_i 势向量依赖于一个变量 x 并有形式 $A_i(x-ct), i=0,1,2,3$. 如果 $A_i(x)$ 为光滑函数, 当所有 x 的模有公共的常数上界时, 则电磁场的所有代数不变量 (特征值) 恒等于零 (证明!). 这是电磁波沿 x 轴传播的情形.

4. 引力场方程

设在闵可夫斯基空间 M^4 中度量为 g_{ij}, Γ^i_{jk} 为与此度量相容的对称联络. 以 x^0,x^1,x^2,x^3 表示 M^4 中的坐标. 在物理学中,M^4 中的长度元的平方 dl^2 等同于引力场, $dl^2=g_{ij}dx^idx^j, g_{ij}=g_{ji}$. 设 $g=\det(g_{ij})$. 在惯性参照系中, 当利用欧氏空间坐标 $x^1=x, x^2=y, x^3=z$ 及时间坐标 $x^0=ct$ 时, 我们得到了 $dl^2=(dx^0)^2-(dx^1)^2-(dx^2)^2-(dx^3)^2$ (我们记得这种坐标叫做伪欧氏的). 引进了这种坐标的这个时空空间被称做平坦空间.

一般讲, 我们将在 M^4 中考虑变化的伪黎曼度量 (常称这样的时空空间 M^4 为弯曲空间, 以和平坦的加以区别). 当然, 在每个特定点 $x_0\in M^4$ 上可以化 $g_{ij}(x_0)$ 为对角形, 这是经过在点 x_0 的某个邻域内相应的坐标变换完成的. 在化为对角形后, 度量张量的矩阵 (g_{ij}) 有一个正的, 三个负的特征值, 因此 $g=\det(g_{ij})<0$.

设 $d\Omega=\sqrt{-g}d^4x$ 是标准的 4 元体积形式. 以 R^i_{jkl} 表示黎曼曲率张量, 它由与 g_{ij} 相容的仿射联络 Γ^i_{jk} 构造的. 我们记得在 §30 中, 当降低张量 R^i_{jkl} 的上指标后得到下面通过 g_{ij} 和度量张量的导数来表示 R_{ijkl} 的显式公式:

$$\begin{aligned}R_{iklm}&=\frac{1}{2}\left(\frac{\partial^2 g_{im}}{\partial x^k\partial x^l}+\frac{\partial^2 g_{kl}}{\partial x^i\partial x^m}-\frac{\partial^2 g_{il}}{\partial x^k\partial x^m}-\frac{\partial^2 g_{km}}{\partial x^i\partial x^l}\right)+\\&\quad g_{np}(\Gamma^n_{kl}\Gamma^p_{im}-\Gamma^n_{km}\Gamma^p_{il}).\end{aligned}$$

对于里奇张量 $R_{ik}=R^q_{iqk}=g^{lm}R_{limk}$, 我们容易得到下面的表达式

$$R_{ik}=\frac{\partial \Gamma^l_{ik}}{\partial x^l}-\frac{\partial \Gamma^l_{il}}{\partial x^k}+\Gamma^l_{ik}\Gamma^m_{lm}-\Gamma^m_{il}\Gamma^l_{km}$$

(请验证!). 我们也已知道，标量曲率 R 就是下面的不变量:

$$R = g^{ik}R_{ik} = g^{il}g^{km}R_{iklm}.$$

对于引力场的爱因斯坦方程可以作为这个场的希尔伯特作用的变分得到的欧拉 – 拉格朗日方程. 我们记这个作用的泛函为 S_g. 假设 $S_g = \int R d\Omega$, 其中的积分取在整个三维空间 (x^1, x^2, x^3), 以及时间坐标 x^0 在区间 $x_1^0 \leqslant x^0 \leqslant x_2^0$ 上.

定理 37.2. 成立恒等式

$$\frac{\delta \int R\sqrt{|g|}d^4x}{\delta g^{ij}} = (R_{ij} - \frac{1}{2}Rg_{ij})\sqrt{|g|},$$

从而

$$\delta S_g = \int (R_{ik} - \frac{1}{2}Rg_{ik})\delta g^{ik}\sqrt{|g|}d^4x.$$

证明 我们对 S_g 的表达式可以进行某种简化，以消去出现在不变量 R 中的二阶导数 $\partial^2 g_{ik}/\partial x^p \partial x^q$. 在这个简化 (我们现在就进行) 之后，被积表达式就只包含了张量 g_{ik} 和克氏符号 Γ^i_{jk}.

我们有

$$R\sqrt{-g} = \sqrt{-g}g^{ik}R_{ik} = \sqrt{-g}\left(g^{ik}\frac{\partial \Gamma^l_{ik}}{\partial x^l} - g^{ik}\frac{\partial \Gamma^l_{il}}{\partial x^k} + g^{ik}\Gamma^l_{ik}\Gamma^m_{lm} - g^{ik}\Gamma^m_{il}\Gamma^l_{km}\right).$$

改变前两项:

$$\sqrt{-g}g^{ik}\frac{\partial \Gamma^l_{ik}}{\partial x^l} = \frac{\partial}{\partial x^l}(\sqrt{-g}g^{ik}\Gamma^l_{ik}) - \Gamma^l_{ik}\frac{\partial}{\partial x^l}(\sqrt{-g}g^{ik}),$$

$$\sqrt{-g}g^{ik}\frac{\partial \Gamma^l_{il}}{\partial x^k} = \frac{\partial}{\partial x^k}(\sqrt{-g}g^{ik}\Gamma^l_{il}) - \Gamma^l_{il}\frac{\partial}{\partial x^k}(\sqrt{-g}g^{ik}).$$

可以去掉全导数 (准确地说，是发散量) $\frac{\partial}{\partial x^l}(\sqrt{-g}g^{ik}\Gamma^l_{ik})$ 和 $\frac{\partial}{\partial x^k}(\sqrt{-g}g^{ik}\Gamma^l_{il})$ 而不会影响变差 δS_g. 关键在于积分

$$\int_D \frac{\partial}{\partial x^l}(\sqrt{-g}g^{ik}\Gamma^l_{ik})d^4x, \quad \int_D \frac{\partial}{\partial x^k}(\sqrt{-g}g^{ik}\Gamma^l_{il})d^4x$$

的变差等于零. 事实上，利用斯托克斯公式，这些积分可以变为在区域 D 的边界 ∂D 上的积分. 因为在作 S_g 的变分时，我们应该对 g_{ik} 变差 (把它们看作独立的变量)，故而变差 δg_{ik} 应该在 ∂D 上等于零 (参照上面的欧拉 – 拉格朗日方程的结论). 因此 $\delta(\int_{\partial D}) = 0$. 于是可以认为

$$\delta(\int R d\Omega) = \delta(\int G\sqrt{-g}d^4x),$$

其中

$$G\sqrt{-g} = \Gamma^l_{il}\frac{\partial}{\partial x^k}(\sqrt{-g}g^{ik}) - \Gamma^l_{ik}\frac{\partial}{\partial x^l}(\sqrt{-g}g^{ik}) - (\Gamma^m_{il}\Gamma^l_{km} - \Gamma^l_{ik}\Gamma^m_{lm})g^{ik}\sqrt{-g}. \tag{14}$$

由于联络 Γ^i_{jk} 与度量相容, 则成立恒等式 (参看习题 29.11)

$$g^{kl}\Gamma^i_{kl} = -\frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^k}(\sqrt{-g}g^{ik})$$

(请验证!). 以 $\sqrt{-g}$ 去除 (14), 我们得到

$$G = -\Gamma^l_{il}\Gamma^i_{kp}g^{kp} - \Gamma^l_{ik}\frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^l}(\sqrt{-g}g^{ik}) - (\Gamma^m_{il}\Gamma^l_{km} - \Gamma^l_{ik}\Gamma^m_{lm})g^{ik}.$$

变换第二项:

$$\begin{aligned}-\Gamma^l_{ik}\frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^l}(\sqrt{-g}g^{ik}) &= \Gamma^l_{ik}\frac{1}{\sqrt{-g}}\frac{1}{2\sqrt{-g}}\frac{\partial g}{\partial x^l}g^{ik} - \Gamma^l_{ik}\frac{\partial g^{ik}}{\partial x^l}\\ &= -\Gamma^l_{ik}g^{ik}\Gamma^p_{lp} - \Gamma^l_{ik}\frac{\partial g^{ik}}{\partial x^l}.\end{aligned}$$

在这里我们利用了关系式 $\Gamma^i_{ki} = \dfrac{1}{2g}\dfrac{\partial g}{\partial x^k}$. 我们还知道, $\dfrac{\partial g^{ik}}{\partial x^l} = -\Gamma^i_{ml}g^{mk} - \Gamma^k_{ml}g^{im}$. 由此有

$$\begin{aligned}-\Gamma^l_{ik}\Gamma^p_{lp}g^{ik} - \Gamma^l_{ik}\frac{\partial g^{ik}}{\partial x^l} &= -\Gamma^l_{ik}\Gamma^p_{lp}g^{ik} + \Gamma^l_{ik}\Gamma^i_{ml}g^{mk} + \Gamma^l_{ik}\Gamma^k_{ml}g^{im}\\ &= -\Gamma^l_{ik}\Gamma^m_{lm}g^{ik} + \Gamma^l_{ik}\Gamma^i_{ml}g^{mk} + \Gamma^i_{ml}\Gamma^l_{ki}g^{km}\\ &= -\Gamma^l_{ik}\Gamma^m_{lm}g^{ik} + 2\Gamma^l_{ik}\Gamma^i_{ml}g^{mk}.\end{aligned}$$

由此有

$$\begin{aligned}G &= 2\Gamma^l_{ik}\Gamma^i_{ml}g^{mk} - \Gamma^l_{ik}\Gamma^m_{lm}g^{ik} - \Gamma^m_{im}\Gamma^i_{kl}g^{kl} - g^{ik}(\Gamma^m_{il}\Gamma^l_{km} - \Gamma^l_{ik}\Gamma^m_{lm})\\ &= g^{ik}(2\Gamma^l_{mk}\Gamma^m_{il} - \Gamma^l_{ik}\Gamma^m_{lm} - \Gamma^m_{lm}\Gamma^l_{ki}) - g^{ik}(\Gamma^m_{il}\Gamma^l_{km} - \Gamma^l_{ik}\Gamma^m_{lm})\\ &= 2g^{ik}(\Gamma^l_{mk}\Gamma^m_{il} - \Gamma^l_{ik}\Gamma^m_{lm}) - g^{ik}(\Gamma^m_{il}\Gamma^l_{km} - \Gamma^l_{ik}\Gamma^m_{lm})\\ &= g^{ik}(\Gamma^m_{il}\Gamma^l_{mk} - \Gamma^l_{ik}\Gamma^m_{lm}).\end{aligned}$$

因此, 我们已证明了下面的命题 (显然, 这个证明对任意的维数 n 都成立):

引理 37.3.

$$\delta\left(\int R\sqrt{|g|}d^n x\right) = \delta\left(\int g^{ik}\left(\Gamma^m_{il}\Gamma^l_{mk} - \Gamma^l_{ik}\Gamma^m_{lm}\right)\sqrt{|g|}d^n x\right).$$

我们再次指出, 引力场由度量张量 g_{ij} 定义, 因此在进行对作用的变差时应该让分量 g_{ij} 接受变分, 而且在此时 g_{ij} 被看为独立的变量.

我们来求 δS_g (只对引力场即度量进行变分). 我们有

$$\begin{aligned}\delta\int R\sqrt{-g}d^4x &= \delta\int g^{ik}R_{ik}\sqrt{-g}d^4x\\ &= \int(R_{ik}\sqrt{-g}\delta g^{ik}+R_{ik}g^{ik}\delta(\sqrt{-g})+g^{ik}\sqrt{-g}\delta R_{ik})d^4x.\end{aligned}$$

求 δg^{ik}. 如果 Δ^{ik} 为 g_{ik} 在矩阵 (g_{jl}) 中的余子式, 则 $g=\sum_i g_{ik}\Delta^{ik}$ (按 g 的定义), 其中取和仅对 i 进行, 而 k 是固定的. 明显地, 因为每个分量 g_{ik} 的微分 δg_{ik} 必须乘以 (在约化相似项时) 这个分量在 g 的表达式中的系数, 即 Δ^{ik}, 则 $\delta g=(\delta g_{ik})\Delta^{ik}$ (对 i,k 取和). 因为 $g^{ik}=\dfrac{\Delta^{ik}}{g}$, 故 $\Delta^{ik}=g\cdot g^{ik}$, 从而 $\delta g=gg^{ik}\delta g_{ik}$. 因为 $g^{ik}g_{ik}=\delta_i^i=4$, 故 $(\delta g^{ik})g_{ik}+g^{ik}(\delta g_{ik})=0$, 即 $\delta g=-gg_{ik}\delta g^{ik}$. 于是,

$$\delta(\sqrt{-g})=\frac{-1}{2\sqrt{-g}}\delta g=\frac{1}{2\sqrt{-g}}g\cdot g_{ik}\delta g^{ik}=-\frac{1}{2}\sqrt{-g}g_{ik}\delta g^{ik}.$$

由此得到

$$\delta\int R\sqrt{-g}d^4x=\int(R_{ik}-\frac{1}{2}Rg_{ik})\delta g^{ik}\sqrt{-g}d^4x+\int g^{ik}(\delta R_{ik})\sqrt{-g}d^4x.$$

现在来求 δR_{ik}. 我们先要注意到, 变差 $\delta\Gamma^i_{jk}$ 是一个张量 (记住, 克里斯托费尔符号 Γ^i_{jk} 不构成向量). 事实上, 按沿向量场的共变微分的运算 $\nabla_\xi=\xi^i\nabla_i$ 的定义, 我们得到 $\nabla_{\partial_i}(\partial_j)=\Gamma^k_{ij}\partial_k$, 其中 ∂_i 为标准坐标向量场 $\partial_i=\dfrac{\partial}{\partial x^i}$. 因此,$\Gamma^i_{jk}$ 为沿坐标曲线平行移动的向量展式中的系数. 于是, 表达式 $\delta\Gamma^k_{ij}\partial_k$ 是两个向量的差, 这两个向量是以两种方式从点 Q 到点 $\tilde{Q}$ 的平行移动所得到的: 一个是用原来的联络 Γ^k_{ij}, 另一个是用变差的联络 $\Gamma^k_{ij}+\delta\Gamma^k_{ij}$, 因为在同一个点 $\tilde{Q}$ 上计算的两个向量差仍是一个向量, 故量 $\delta\Gamma^k_{ij}$ 是个张量.

为计算变差 δR_{ik}, 我们固定一个任意点, 并在这个点的某个邻域内引进在此点的惯性坐标系 (参看习题 29.4). 这表明在此点有 $\Gamma^k_{ij}=0$, 因为在这个点有 $\dfrac{\partial}{\partial x^j}(g^{ik})=0$. 我们有

$$\begin{aligned}g^{ik}\delta R_{ik} &= \delta\left(\frac{\partial\Gamma^l_{ik}}{\partial x^l}-\frac{\partial\Gamma^l_{il}}{\partial x^k}+\Gamma^l_{ik}\Gamma^m_{lm}-\Gamma^m_{il}\Gamma^l_{km}\right)g^{ik}\\ &= g^{ik}\left(\frac{\partial}{\partial x^l}(\delta\Gamma^l_{ik})-\frac{\partial}{\partial x^k}(\delta\Gamma^l_{il})\right)\\ &= g^{ik}\frac{\partial}{\partial x^l}(\delta\Gamma^l_{ik})-g^{il}\frac{\partial}{\partial x^l}(\delta\Gamma^k_{ik})\\ &= \frac{\partial}{\partial x^l}(g^{ik}\delta\Gamma^l_{ik}-g^{il}\delta\Gamma^k_{ik})=\frac{\partial W^l}{\partial x^l},\end{aligned}$$

$$W^l = g^{ik}\delta\Gamma^l_{ik} - g^{il}\delta\Gamma^k_{ik}.$$

因为 $\delta\Gamma^i_{jk}$ 为张量, 故量 W^l 也构成了张量 (向量), 从而它的发散量 $\dfrac{\partial W^l}{\partial x^l}$ 在我们这个特殊坐标系 (在所选的点上为惯性的) 下, 当转变到任意其他曲线坐标时不变. 这时因为量 $\dfrac{\partial}{\partial x^i}(T^i)$ 对于任意坐标变换不具张量特性, 我们当然应该运用不变性的定义 $\operatorname{div} T = \nabla_i T^i$. 我们记得 (见 §29.3) 在关于黎曼对称联络的任意坐标中,div T 的显式表达为

$$\nabla_i T^i = \frac{\partial T^i}{\partial x^i} + T^l\frac{\partial}{\partial x^l}(\ln\sqrt{-g}) = \frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^l}(\sqrt{-g}T^l).$$

我们得到了 $g^{ik}\delta R_{ik} = \dfrac{1}{\sqrt{-g}}\dfrac{\partial}{\partial x^l}(\sqrt{-g}W^l)$. 于是积分 $\displaystyle\int g^{ik}\delta R_{ik}\sqrt{-g}d^4x$ 变换为 $\displaystyle\int\frac{\partial}{\partial x^l}(\sqrt{-g}W^l)d^4x$. 由斯托克斯公式, 这个积分变成了在四维区域的边界 ∂D 上对 W^l 的积分. 因为在 ∂D 上场的变差等于零, 故此积分也等于零. 最后我们得到

$$\delta S_g = \int(R_{ik} - \frac{1}{2}Rg_{ik})\delta g^{ik}\sqrt{-g}d^4x.$$

定理 37.2 得证. □

通常作用 S_g 写成具系数 λ 的特殊形式, 其中 $\lambda = \dfrac{c^3}{16\pi G}$, 其中 c 为光速,G 为所谓的 "引力常数". 于是

$$\delta S_g = \frac{c^3}{16\pi G}\int(R_{ik} - \frac{1}{2}Rg_{ik})\delta g^{ik}\sqrt{-g}d^4x.$$

考虑连同物质性在内的总作用 $S_g + S_m = S_{\Pi}$. 我们现在来考虑变差 δS_m. 通常作用 S_m 被写成 (在曲线坐标下) 形式 $\dfrac{1}{c}\displaystyle\int\Lambda\sqrt{-g}d^4x$, 其中 Λ 是由物质的性质定义, 它依赖于由物质定义的度量和场.(回忆: 积分是取在整个三维空间 x^1, x^2, x^3 上, 而在时间 x^0 上是取在两个固定的时刻之间的, 就是说是在无限的四维区域上, 它被包含在两个超曲面 $x^0 = x^0_1$ 和 $x^0 = x^0_2$ 之间;x^0_1, x^0_2 是某两个常数.) 因此

$$\frac{1}{\lambda\sqrt{-g}}\frac{\delta S_g}{\delta g^{ik}} = R_{ik} - \frac{1}{2}Rg_{ik} - \frac{8\pi G}{c^3}\frac{\delta S_m}{\delta g^{ik}}\frac{2}{c\sqrt{-g}}. \tag{15}$$

如果已知 $-\dfrac{2}{c\sqrt{-g}}\dfrac{\delta S_m}{\delta g^{ik}}$, 那么这便是引力场方程. 我们注意到, 这些方程是非线性的, 因此两个 (对场 g_{ij}) 解的和不必是一个解.

在相对论中设定量 $-\dfrac{2}{c\sqrt{-g}}\dfrac{\delta S_m}{\delta g^{ik}} = T_{ik}$ 为系统的能量 — 动量张量. 对物理上重

要的场 (例如电磁场) 而言, 表达式

$$-\frac{2}{c\sqrt{-g}}\frac{\delta S_m}{\delta g^{ik}}\bigg|_{g^{ik}=\text{闵可夫斯基的}g^{ik}}$$

同于原先引进的对称的能量 - 动量张量. 事实上,

$$\begin{aligned} cS_m &= \int \Lambda\sqrt{-g}d^4x = \lambda\int F_{ik}F^{ik}\sqrt{-g}d^4x \\ &= \lambda\int F_{ik}g^{ip}g^{kj}F_{pj}\sqrt{-g}d^4x = \int \Lambda(g_{ab};F_{ik})\sqrt{-g}d^4x. \end{aligned}$$

显然, 我们有 (请验证!)

$$-\frac{2}{c}\delta S_m = \int \frac{1}{4\pi c}(-F_i^lF_{lk} + \frac{1}{4}g_{ik}F_{lm}F^{lm})\delta g^{ik}\sqrt{-g}d^4x.$$

被积表达式与表达式 (13) 相同.

在所谓的 "空" 空间情形 (即不存在物质) 有 $T_{ik}=0$, 于是引力场方程有形式 $R_{ik}=0$. 事实上, 由 (15) 可以得到 $g^{ij}R_{ik}-\frac{1}{2}R\delta_k^j=\frac{8\pi G}{c^4}g^{ij}T_{ik}$, 即 $R_k^j-\frac{1}{2}R\delta_k^j=\frac{8\pi G}{c^4}T_k^j$. 去掉 j,k, 我们得到 $R-2R=\frac{8\pi G}{c^4}T$, 其中 $T=T_j^j$, 即 $R=-\frac{8\pi G}{c^4}T$, 从而我们的场具有形式 $R_{ik}=\frac{8\pi G}{c^4}(T_{ik}-\frac{1}{2}g_{ik}T)$. 如果 $T_{ik}=0$, 从而 $R_{ik}=0$.

由方程 $R_{ik}=0$ 即对此空间的里奇张量等于零, 绝不会推出 "空" 空间是平坦的结论. 时空空间为平坦仅当整个黎曼曲率 R_{jkl}^i 等于零. 如果时空空间是三维的, 则 "空" 空间必定平坦, 这是因为在这种情形下, 黎曼张量可以通过里奇张量表出 (公式 (30.18)). 在二维情形, 由计算得到

定理 37.3. 在度量 g_{ij} 的局部变换下, 量 $S[g]=\int KdS$ 不变, 其中 $dS=\sqrt{g}dx^1\wedge dx^2, K$ 为高斯曲率, 而 $i,j=1,2$.

证明 在二维情形我们有 $K=R/2, R_{ij}=\frac{1}{2}Rg_{ij}=Kg_{ij}$(参看定理 30.4). 泛函 S 的变差等于 (见前面)

$$\delta S=\int\left(R_{ij}-\frac{1}{2}Rg_{ij}\right)\delta g^{ij}dS=0.$$

证完. □

对于 $\mathbb{R}^3$ 中的闭曲面而言, 度量变化的局部性并没有实质性影响, 因而得到了

推论 (高斯 – 博内) 在三维欧氏空间中闭曲面上对高斯曲率的积分, 在曲面的光滑形变下不变

$$\int KdS=\text{常数}.$$

这个常数的数值将在第二卷中得出.

5. 皂膜

考虑光滑的超曲面 $V^{n-1}\subset\mathbb{R}^n$, 譬如, 它由图像形式给出:$x^n=f(x^1,\cdots,x^{n-1})$. 设函数 f 的定义域为 $\mathbb{R}^{n-1}$ 中的有界区域 D; 考虑面积的泛函 $S[f]$, 它定义于所有这些函数 f 的空间上

$$S[f]=\int\limits_D\sqrt{|\det\ A|}d^{n-1}x.$$

这里的 $A=(g_{ij}(x)),x\in D$ 为在曲面 V^{n-1} 上所诱导出的黎曼度量, 而 $d^{n-1}x=dx^1\wedge dx^2\wedge\cdots\wedge dx^{n-1}$, 其中 $x^1,\cdots,x^{n-1}$ 为 D 上的欧氏坐标. 拉格朗日 $\sqrt{|\det\ A|}$ 可以通过函数 f 写成显式表达式.

设 $d\tau^{n-1}$ 为 V^{n-1} 上的 $(n-1)$ 维体积, 于是 $S[f]=\int\limits_D d\tau^{n-1}$. 设 $P\in V^{n-1},n(P)$ 为 V^{n-1} 在点 P 的单位法向量,$\alpha(P)$ 为 $n(P)$ 和 $e_n=(0,\cdots,0,1)$ 之间的夹角 (图 38); 于是

$$S[f]=\int\limits_D d\tau^{n-1}=\int\limits_D\frac{d^{n-1}x}{\cos\ \alpha(P)}.$$

另外, 因为 $n(P)$ 由下面公式给出 (参看 §8.3 中对于 $n=3$ 时的公式):

$$n(P)=\frac{1}{\sqrt{1+\sum\limits_{i=1}^{n-1}(f_{x^i})^2}}(-f_{x^1},\cdots,-f_{x^{n-1}}),$$

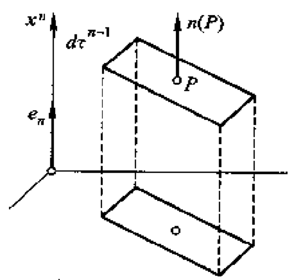


图 38

于是

$$\cos\ \alpha(P)=\langle e_n,n(P)\rangle=\frac{1}{\sqrt{1+\sum\limits_{i=1}^{n-1}(f_{x^i})^2}}.$$

因此, $S[f]=\int\limits_D\sqrt{1+\sum\limits_{i=1}^{n-1}(f_{x^i})^2}dx^1\wedge\cdots\wedge dx^{n-1}$.

考虑对面积泛函 $S[f]$ 的极值曲面 V^{n-1} (即极值函数 $x^n=f(x),x\in D$ 的图像). 欧拉－拉格朗日方程 (3)(这里的 $k=1$) 的形式为

$$\sum_{i=1}^{n-1}\frac{\partial}{\partial x_i}\left(\frac{f_{x^i}}{\sqrt{1+\sum\limits_{j=1}^{n-1}(f_{x^j})^2}}\right)=0.\tag{16}$$

定义 37.5. 称面积泛函 S 的极值曲面为极小曲面.

注　例如在 $\mathbb{R}^3$ 中设计的极小曲面的模型是紧绷在闭金属丝圈上的肥皂膜 (无重力的情形).

对于由 $z=f(x,y)$ 的图像给出的在 $\mathbb{R}^3(x,y,z)$ 中的二维极小曲面, 方程 (16) 得到的形式为

$$(1+f_x^2)f_{yy}-2f_{xy}f_xf_y+(1+f_y^2)f_{xx}=0$$

(请验证!).

极小曲面 V^{n-1} 的方程允许用在 $\mathbb{R}^3$ 中此曲面嵌入的局部不变量去描述它.

定理 37.4. 设 $V^{n-1}\subset\mathbb{R}^n$ 为光滑超曲面. 其平均曲率 H 恒等于零当且仅当 V^{n-1} 在它的每点的某个邻域中表示为面积泛函的极值函数的图像 (即为极小曲面方程的解).

因此, 条件 $H\equiv 0$ 是曲面 $V^{n-1}\subset\mathbb{R}^n$ 的极小性条件.

证明　定理的证明归结为对于图像 $x^n=f(x)$ 的平均曲率 $H=\mathrm{Tr}(A^{-1}Q)$ 的直接计算, 其中 A,Q 分别为第一和第二基本型的矩阵 (H 的这种定义是定义 8.3 在 n 维情形中的直接推广), 并且验证方程 $H=0$ 等同于欧拉 - 拉格朗日方程 (16).　□

我们不在这里进行一般情形下的计算, 而仅仅考虑特殊的情形:$\mathbb{R}^3$ 中的二维极小曲面.

我们以半径向量 $r=r(u,v)$ 给出曲面 $V^2\subset\mathbb{R}^3$(至少是局部地). 于是 $S[r]=\int\limits_{D(u,v)}\sqrt{EG-F^2}dudv$, 其中 $A=\begin{pmatrix}E & F\\ F & G\end{pmatrix}$ 为第一基本型的矩阵:

$$E=\langle r_u,r_u\rangle,\quad F=\langle r_u,r_v\rangle,\quad G=\langle r_v,r_v\rangle.$$

平均曲率 H 的形式为

$$H=\mathrm{Tr}(A^{-1}Q)=\frac{1}{EG-F^2}(GL-2FM+EN),$$

其中 $Q=\begin{pmatrix}L & M\\ M & N\end{pmatrix}$ 为第二基本型的矩阵:

$$L=\langle r_{uu},n\rangle,\quad M=\langle r_{uv},n\rangle,\quad N=\langle r_{vv},n\rangle,$$

这里的 n 是曲面的单位法向量.

在 V^2 上选取 (局部) 共形 (等温) 坐标. 对于实解析度量 A 这样的坐标的存在性早在 §13 中证明过. 为简明起见, 我们假定 u,v 已经是共形坐标了.

在共形坐标下 $F=0,E=G$, 因而

$$S[r]=\iint\limits_D\sqrt{\langle r_u,r_u\rangle\langle r_v,r_v\rangle}dudv=\iint\limits_D\sqrt{(x_u^2+y_u^2+z_u^2)(x_v^2+y_v^2+z_v^2)}dudv.$$

另外, $H = \frac{1}{E}(L+N) = \frac{1}{E}\langle r_{uu}+r_{vv}, n\rangle = \frac{1}{E}\langle \nabla r, n\rangle$, 其中 ∇ 为拉普拉斯算子.

我们考虑在坐标 u, v 下 $S[r]$ 的欧拉 – 拉格朗日方程. 这表明, 将欧拉 – 拉格朗日方程以共形坐标描述的可能性来自于下面的事实, 即当用扰动 η 对泛函 $S[r]$ 进行变分时应可以假定所有函数 $r(u,v)+\varepsilon\eta(u,v)$ 均参照于共形坐标. 为此, 只要在每个曲面 $r(u,v)+\varepsilon\eta(u,v)$ 上引进共形坐标 $(u_\varepsilon, v_\varepsilon)$ (关键是坐标 u, v 在扰动了的曲面 $r+\varepsilon\eta$ 上一般已不是共形的了). 假定坐标 $(u_\varepsilon, v_\varepsilon)$ 光滑依赖于 ε.

欧拉 – 拉普拉斯方程具形式

$$\left.\begin{aligned} &\frac{\partial}{\partial u}(x_u)+\frac{\partial}{\partial v}(x_v)=0,\\ &\frac{\partial}{\partial u}(y_u)+\frac{\partial}{\partial v}(y_v)=0,\\ &\frac{\partial}{\partial u}(z_u)+\frac{\partial}{\partial v}(z_v)=0. \end{aligned}\right\},\text{即}\Delta r=0$$

称满足方程 $\Delta r = 0$ 的半径向量 r 为调和的. 因此, 在共形坐标下 "半径向量的极小性"(即它对面积泛函的极小性) 意味着它的调和性.

因此, 由于 $\Delta r = 0$, 则 $H = \frac{1}{E}(\Delta r, n) = 0$, 从而证明了我们感兴趣的命题的一个方面.

反之, 设 $H \equiv 0$, 我们应该证明 $\Delta r = 0$ (在共形坐标中). 因为 $\langle \Delta r, n\rangle = 0$, 故只要验证两个等式:$\langle \Delta r, r_u\rangle = 0, \langle \Delta r, r_v\rangle = 0$. 由此将推出 $\Delta r = 0$. 事实上, 向量 n, r_u, r_v 构成在曲面 $r(u,v)$ 的每个正则点上构成一个标架 (根据曲面的定义). 由于坐标的选取, 我们有 $E = G$ 和 $F = 0$, 即 $\langle r_u, r_u\rangle = \langle r_v, r_v\rangle$ 以及 $\langle r_u, r_v\rangle = 0$. 按 u 和 v 作微分, 得到

$$\begin{aligned} \langle r_{uu}, r_u\rangle &= \langle r_{uv}, r_v\rangle,\\ \langle r_{uv}, r_u\rangle &= \langle r_{vv}, r_v\rangle,\\ \langle r_{uu}, r_v\rangle + \langle r_u, r_{uv}\rangle &= 0,\\ \langle r_{uv}, r_v\rangle + \langle r_u, r_{vv}\rangle &= 0. \end{aligned} \tag{17}$$

我们应验证恒等式

$$\begin{aligned} \langle r_{uu}, r_u\rangle + \langle r_{vv}, r_u\rangle &= 0,\\ \langle r_{uu}, r_v\rangle + \langle r_{vv}, r_v\rangle &= 0. \end{aligned}$$

最后面的这两个方程显然可由 (17) 得到. 从而证明了

命题 37.2. 二维曲面可由极小半径向量描述当且仅当 $H \equiv 0$. 在共形坐标中极小半径向量成为调和向量.

注 论及半径向量 $r(u,v)$ 的调和性可能仅仅与某种坐标系有关 (在上述给出的情形中这个坐标是共形的). 在坐标变换下, 调和性一般说来是不能保持的.

极小曲面 $V^2 \subset \mathbb{R}^3$ 的结构十分复杂. 例如, 如果固定边缘的闭曲线 $S^1 \subset \mathbb{R}^3$, 则一般情形, 在其上可以张成许多种 "皂膜" (即对于微分方程 $H \equiv 0$ 或 $\Delta r \equiv 0$ 的解没有唯一性定理.) 参看图 39,40 上的例子.

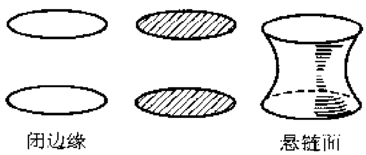


图 39

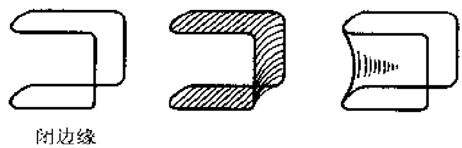


图 40

方程 $H = 0, \Delta r = 0$ 的解可能有奇点. 例如 (三叶默比乌斯曲面) 参看图 41,42.

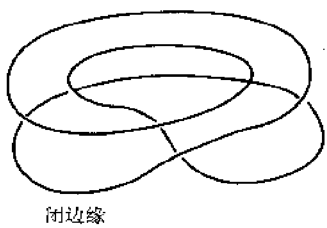


图 41

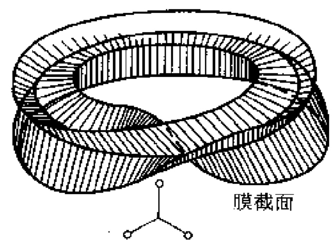


图 42

调和的半径向量是对一个高维 (二维) 泛函的欧拉 – 拉格朗日方程的解.

狄利克雷泛函　考虑二维半径向量 $r(u, v)$ (坐标 u, v 为任意). 称下面的泛函

$$D[r] = \int\limits_{D(u,v)} \frac{E+G}{2} du dv$$

为狄利克雷泛函, 其中 E, G 为曲面 $r(u, v)$ 的第一基本型的系数.

这里

$$L(r_u, r_v) = \frac{1}{2}(x_u^2 + y_u^2 + z_u^2 + x_v^2 + y_v^2 + z_v^2),$$

故而欧拉 – 拉格朗日方程 (向量形式) 的形式为 $\Delta r = 0$. 其解便是调和半径向量.

因为

$$\frac{E+G}{2} \geqslant \sqrt{EG - F^2},$$

故

$$D[r] \geqslant S[r]$$

对任意逐段光滑的半径向量 $r(u, v)$ 成立, 而等号成立当且仅当 $E = G, F = 0$, 即在共形坐标系中. 于是, 任意 $D(r)$ 的极值面, 其坐标 u, v 为共形时则也是 $S[r]$ 的极

值; 反之则不成立. 为了得到泛函 $S[r]$ 的所有极值面, 应该考虑所有泛函 $D[r]$ 的极值 (即调和半径向量), 从它们中只选出那些对它们而言 u, v 是共形坐标, 然后对 u, v 进行任意正则坐标变换.

对坐标 u, v 不是共形的调和半径向量 $r(u, v)$, 将不能被用来描述极小曲面. 例如,$r(u, v) = (u, v, \mathrm{Re}\ f(u + iv))$ 为非线性复解析函数 $f(u + iv)$ 的实部 (或虚部) 的图像.

泛函 $D[r]$ 和 $S[r]$ 之间的联系, 在许多方面类似于长度泛函 $l_a^b[\gamma]$ 和在道路 γ 上的作用 $S_a^b[\gamma] = \int_a^b |\dot\gamma|^2 dt$ 之间的联系. 显见,$(l_a^b[\gamma])^2 \leqslant (b - a) S_a^b[\gamma]$, 而等号成立当且仅当在轨线 $\gamma(t)$ 上参数 t 与弧长成比例 (如果参数是自然的, 这样的极值 $l_a^b[\gamma]$ 为测地线). 这种情形与下面情形相关联, 即泛函 $l_a^b[\gamma]$ 和 $S[r]$ 对于变量变换不变, 而泛函 $S_a^b[\gamma]$ 和 $D[r]$ 则不是不变的.

6. 薄板的平衡方程

考虑一个进行形变的物体的一个特殊情形: 一个弯曲薄片的平衡. 我们假定这个片是薄的 (即其厚度与其他两个方向的尺度相比是小的). 假定在无形变情形下这个片是个平面. 还将假定形变是小的, 即薄板上点的位移相对于厚度而言是小的. 对薄板的自由 (弹性) 能量变分的欧拉 – 拉格朗日方程就是薄板的平衡方程.

在弯曲薄板时, 在一些点产生了伸张, 而在另一些点产生了压缩. 在凸的一边进行了伸张, 而在凹的一边进行了压缩. 按深入厚度方向的程度, 伸张和压缩在减少; 伸张的区域和压缩的区域彼此被一个所谓的 "中性曲面"隔开, 在此曲面不存在伸或缩. 这个曲面位于薄板厚度的中间.

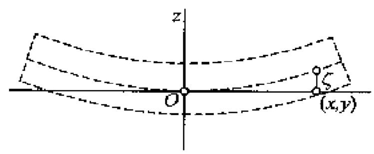


图 43

我们引进笛卡儿坐标系 x, y, z, 其原点 O 在中性曲面上,Oz 轴正交于此曲面. 平面 (x, y) 与未形变时的薄板相重合. 以 $\zeta(x, y)$ 表示在弯曲时中性曲面的点的竖直位移 (图 43).

设 h 为板的厚度, 于是弯曲薄板的总自由能量 F 由下面的公式计算:

$$F = \frac{Eh^3}{24(1-\sigma^2)} \iint \left\{ \left(\frac{\partial^2\zeta}{\partial x^2} + \frac{\partial^2\zeta}{\partial y^2} \right)^2 + 2(1-\sigma) \left[\left(\frac{\partial^2\zeta}{\partial x \partial y} \right)^2 - \frac{\partial^2\zeta}{\partial x^2} \frac{\partial^2\zeta}{\partial y^2} \right] \right\} dxdy$$

其中的积分在位移函数 $\zeta = \zeta(x, y)$ 的整个定义域上进行,E 为伸缩模量 (杨氏模量),σ 为泊松系数, 它们的计算公式为

$$E = \frac{qK\mu}{3K + \mu}, \quad \sigma = \frac{1}{2} \frac{3K - 2\mu}{3K + \mu},$$

其中 K 和 μ 分别是总的压缩模数和剪切模数; 常数 K 和 μ 由薄板的组成物质的性质所决定.

以下来推导平衡方程. 我们注意到, 由于是小的形变, 故可设 $dxdy = dS$, 其中 dS 为曲面的面积元 (中性曲面). 我们对能量 F 进行变分. 把 F 表示为两个积分之和:

$$F = \frac{Eh^3}{24(1-\sigma^2)} \left\{ \iint (\Delta\zeta)^2 dS + 2(1-\sigma) \iint \left[\left(\frac{\partial^2 \zeta}{\partial x \partial y} \right)^2 - \frac{\partial^2 \zeta}{\partial x^2} \frac{\partial^2 \zeta}{\partial y^2} \right] dS \right\}$$

(这里的 Δ 是拉普拉斯算子) 从而分别对这些积分进行变差.

$$\begin{aligned} \delta \frac{1}{2} \iint (\Delta\zeta)^2 dS &= \iint (\Delta\zeta)(\Delta\delta\zeta) dS = \iint (\Delta\zeta)(\mathrm{div\ grad}\ \delta\zeta) dS \\ &= \iint \mathrm{div}(\Delta\zeta \mathrm{grad}\ \delta\zeta) dS - \iint \langle \mathrm{grad}\ \Delta\zeta, \mathrm{grad}\ \delta\zeta \rangle dS. \end{aligned}$$

设 $\gamma = \partial D$ 为闭曲线, 它是函数 $\zeta(x,y)$ 定义域 D 的边缘. 例如,γ 是包围这个薄板的闭曲线. 于是积分 $\iint\limits_D \mathrm{div}(\Delta\zeta \mathrm{grad}\ \delta\zeta) dS$ 可以用斯托克斯公式变化为 $\oint\limits_\gamma \Delta\zeta \langle n, \mathrm{grad}\ \delta\zeta \rangle dl$, 其中 dl 是沿 γ 的弧长元,n 为 γ 的外法线. 可清楚看出

$$\oint\limits_\gamma \Delta\zeta \langle n, \mathrm{grad}\ \delta\zeta \rangle dl = \oint\limits_\gamma \Delta\zeta \frac{\partial(\delta\zeta)}{\partial n} dl,$$

其中 $\frac{\partial}{\partial n}$ 为沿闭曲线 γ 的外法线 n 的方向微分. 类似地,

$$\begin{aligned} \iint\limits_D \langle \mathrm{grad}\ \delta\zeta, \mathrm{grad}\ \Delta\zeta \rangle dS &= \iint\limits_D \mathrm{div}((\delta\zeta)\mathrm{grad}\Delta\zeta) - \iint\limits_D (\delta\zeta)\Delta^2\zeta dS \\ &= \oint\limits_\gamma \delta\zeta \langle n, \mathrm{grad}\ \Delta\zeta \rangle dl - \iint\limits_D (\delta\zeta)\Delta^2\zeta dS \\ &= \oint\limits_\gamma \delta\zeta \frac{\partial(\Delta\zeta)}{\partial n} dl - \iint\limits_D (\delta\zeta)\Delta^2\zeta dS. \end{aligned}$$

因此, 我们对 δF 表达式中的第一个积分化为

$$\delta \frac{1}{2} \iint\limits_D (\Delta\zeta)^2 dS = \iint\limits_D (\delta\zeta)\Delta^2\zeta dS - \oint\limits_\gamma \delta\zeta \frac{\partial(\Delta\zeta)}{\partial n} dl + \oint\limits_\gamma \Delta\zeta \frac{\partial(\delta\zeta)}{\partial n} dl.$$

转而考虑 δF 表达式中的第二个积分:

$$\begin{aligned} &\delta \iint\limits_D \left[\left(\frac{\partial^2 \zeta}{\partial x \partial y} \right)^2 - \frac{\partial^2 \zeta}{\partial x^2} \frac{\partial^2 \zeta}{\partial y^2} \right] \delta S \\ &= \iint\limits_D \left[2 \frac{\partial^2 \zeta}{\partial x \partial y} \frac{\partial^2 \delta\zeta}{\partial x \partial y} - \frac{\partial^2 \zeta}{\partial x^2} \frac{\partial^2 \delta\zeta}{\partial y^2} - \frac{\partial^2 \delta\zeta}{\partial x^2} \frac{\partial^2 \zeta}{\partial y^2} \right] dS. \end{aligned}$$

其中被积表达式可以表示为下面的形状:

$$\frac{\partial}{\partial x}\left(\frac{\partial\delta\zeta}{\partial y}\frac{\partial^2\zeta}{\partial x\partial y}-\frac{\partial\delta\zeta}{\partial x}\frac{\partial^2\zeta}{\partial y^2}\right)+\frac{\partial}{\partial y}\left(\frac{\partial\delta\zeta}{\partial x}\frac{\partial^2\zeta}{\partial x\partial y}-\frac{\partial\delta\zeta}{\partial y}\frac{\partial^2\zeta}{\partial x^2}\right)=\operatorname{div} T.$$

由此知

$$\begin{aligned}&\delta\iint_D\left[\left(\frac{\partial^2\zeta}{\partial x\partial y}\right)^2-\frac{\partial^2\zeta}{\partial x^2}\frac{\partial^2\zeta}{\partial y^2}\right]dS\\ =&\oint_\gamma(n,T)dl\\ =&\oint_\gamma\left[\cos\ \theta\left(\frac{\partial\delta\zeta}{\partial y}\frac{\partial^2\zeta}{\partial x\partial y}-\frac{\partial\delta\zeta}{\partial x}\frac{\partial^2\zeta}{\partial y^2}\right)+\sin\ \theta\left(\frac{\partial\delta\zeta}{\partial x}\frac{\partial^2\zeta}{\partial x\partial y}-\frac{\partial\delta\zeta}{\partial y}\frac{\partial^2\zeta}{\partial x^2}\right)\right]dl,\end{aligned}$$

其中 θ 为 Ox 轴和法向量 n 之间的夹角 (图 44). 我们来用 $\dfrac{\partial\delta\zeta}{\partial n}$ 和 $\dfrac{\partial\delta\zeta}{\partial l}$ 来表示 $\dfrac{\partial\delta\zeta}{\partial x}$ 和 $\dfrac{\partial\delta\zeta}{\partial y}$, 其中 $\dfrac{\partial}{\partial n}=\partial_n$. 因为 $\dfrac{\partial}{\partial x}=\cos\ \theta\dfrac{\partial}{\partial n}-\sin\ \theta\dfrac{\partial}{\partial l}$, $\dfrac{\partial}{\partial y}=\sin\ \theta\dfrac{\partial}{\partial n}+\cos\ \theta\dfrac{\partial}{\partial l}$, 于是

$$\begin{aligned}&\delta\iint_D\left[\left(\frac{\partial^2\zeta}{\partial x\partial y}\right)^2-\frac{\partial^2\zeta}{\partial x^2}\frac{\partial^2\zeta}{\partial y^2}\right]dS\\ =&\oint_\gamma\frac{\partial\delta\zeta}{\partial n}\left[2\sin\ \theta\cos\ \theta\frac{\partial^2\zeta}{\partial x\partial y}-\sin^2\theta\frac{\partial^2\zeta}{\partial x^2}-\cos^2\theta\frac{\partial^2\zeta}{\partial y^2}\right]dl+\\ &\oint_\gamma\frac{\partial\delta\zeta}{\partial l}\left[\sin\ \theta\cos\ \theta\left(\frac{\partial^2\zeta}{\partial y^2}-\frac{\partial^2\zeta}{\partial x^2}\right)+(\cos^2\theta-\sin^2\theta)\frac{\partial^2\zeta}{\partial x\partial y}\right]dl.\end{aligned}$$

可对第二个积分进行分部积分. 因为积分是在闭曲线上进行的, 故积分限合为一个点, 因而积分具有形式

$$-\oint_\gamma\delta\zeta\frac{\partial}{\partial l}\left[\sin\ \theta\cos\ \theta\left(\frac{\partial^2\zeta}{\partial y^2}-\frac{\partial^2\zeta}{\partial x^2}\right)+(\cos^2\theta-\sin^2\theta)\frac{\partial^2\zeta}{\partial x\partial y}\right]dl$$

(第二个积分是对于全微分的积分, 故化为零). 最后, 对于自由能量的变分我们有

$$\begin{aligned}\delta F=&\frac{Eh^3}{12(1-\sigma^2)}\left\{\iint_D(\delta\zeta)\Delta^2\zeta dS-\oint_\gamma\delta\zeta\left[\frac{\partial(\Delta\zeta)}{\partial n}+\right.\right.\\ &\left.(1-\sigma)\frac{\partial}{\partial l}\left(\sin\ \theta\cos\ \theta\left(\frac{\partial^2\zeta}{\partial y^2}-\frac{\partial^2\zeta}{\partial x^2}\right)+(\cos^2\theta-\sin^2\theta)\frac{\partial^2\zeta}{\partial x\partial y}\right)\right]dl+\\ &\left.\oint_\gamma\frac{\partial(\delta\zeta)}{\partial n}\left[\Delta\zeta+(1-\sigma)\left(2\sin\ \theta\cos\ \theta\frac{\partial^2\zeta}{\partial x\partial y}-\sin^2\theta\frac{\partial^2\zeta}{\partial x^2}-\cos^2\theta\frac{\partial^2\zeta}{\partial y^2}\right)\right]dl\right\}.\end{aligned}$$

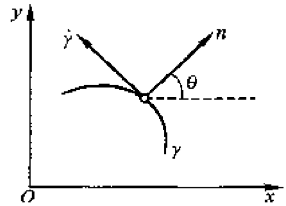


图 44

为了得到薄板的平衡方程 (欧拉 – 拉格朗日方程) 需要使和 $\delta F+\delta U$ 为零, 其中 U 为薄板的势能, 它与外力在它上的作用相关联. 变差 δU 等于在薄板位移时外力的功加上反向符号.

设 $P=P(x,y)$ 为作用在薄板的外力, 它是赋在它的曲面的单位面积上的 (即中性曲面上), 而且是沿其法线方向. 于是外力在薄板位移距离 $\delta\zeta$ 时所做的功等于 $\iint\limits_D P\delta\zeta dS$. 由此, 作为总自由能量的极小性条件 (更准确地说, 极值条件), 我们得到

$$\delta F-\iint\limits_D P\delta\zeta dS=0. \tag{18}$$

在这个关系式中既涉及曲面, 也涉及闭曲线上的积分. 由于变差 $\delta\zeta$ 是任意的, 并且可以有任意小的支集 (特别,$\delta\zeta$ 的支集与闭边缘 γ 不交), 因而分别地在曲面上的和在闭边缘上的积分为零. 曲面积分为

$$\iint\limits_D\left(\frac{Eh^3}{12(1-\sigma)}\Delta^2\zeta-P\right)d\zeta dS=0.$$

由 $\delta\zeta$ 的任意性, 有 $H\Delta^2\zeta-P=0, H=\dfrac{Eh^3}{12(1-\sigma)}$, 称这个量为在弯曲时薄板的刚度 (或柱面刚度); H 由物质的性质决定. 因此, 薄板在外力的弯曲作用下的平衡方程为

$$H\Delta^2\zeta-P=0.$$

这个方程必须满足边界条件, 它由方程 (18) 中的在闭曲线上的积分等于零得到. 通常在此时可选出几个重要的特殊情形.

a) 设薄板边缘 $\gamma=\partial D$ 部分是自由的, 即在其上没有外力作用. 于是沿此边界部分 $\delta\zeta$ 和 $\delta\left(\dfrac{\partial\zeta}{\partial n}\right)$ 为任意, 因此在这些变差中对应于闭曲线积分的那些系数应该为零. 这给出了下面的一组边界条件:

$$-\frac{\partial(\Delta\zeta)}{\partial n}+(1-\sigma)\frac{\partial}{\partial l}\left(\sin\ \theta\cos\ \theta\left(\frac{\partial^2\zeta}{\partial x^2}-\frac{\partial^2\zeta}{\partial y^2}\right)+(\sin^2\theta-\cos^2\theta)\frac{\partial^2\zeta}{\partial x\partial y}\right)=0,$$

$$\Delta\zeta+(1-\sigma)\left(2\sin\ \theta\cos\ \theta\frac{\partial^2\zeta}{\partial x\partial y}-\sin^2\theta\frac{\partial^2\zeta}{\partial x^2}-\cos^2\theta\frac{\partial^2\zeta}{\partial y^2}\right)=0.$$

b) 设薄板的边被刚性固定 (例如被刚性地嵌入到物质之中). 这时板不能进行任何竖直的扰动, 也不能改变板的边缘的方向, 从而 $\delta\zeta\equiv 0,\delta\left(\dfrac{\partial\zeta}{\partial n}\right)=0$ (由此得到, 所

有的闭曲线上积分恒等于零); 边界条件有简单形式:$\zeta = 0, \dfrac{\partial \zeta}{\partial n} = 0$. 第一个方程表明,边界当形变时没有扰动, 而第二个表明边缘的方向在水平形变时保持不变.

习题

37.2. 证明对于泛函 $S[F] = \int F \wedge *F = -\frac{1}{2}\int F_{ik}F^{ik}d^4x$ 在条件 $d(F_{ik}dx^i \wedge dx^k) = 0$ 下的极值麦克斯韦方程成立 (在"空" 空间中). 这里的 F_{ik} 是闵可夫斯基空间 $\mathbb{R}^4_1$ 中的反称张量.

37.3. 证明

a) $P^i = \dfrac{1}{c}\int T^{i0}d^3x$; $d^3x = dx^1 \wedge dx^2 \wedge dx^3$.

称分量为 $\frac{1}{c}T^{10}, \frac{1}{c}T^{20}, \frac{1}{c}T^{30}$ 为系统的动量密度, 而 T^{00} 为能量密度. 如果 V 为超曲面 $x^0 =$ 常数上的有界区域, 于是

b) $\dfrac{\partial}{\partial t}\int\limits_V T^{00}d^3x = -c\oint\limits_{\partial V} T^{0\alpha}d\sigma_\alpha$, 其中 $T^{0\alpha}d\sigma_\alpha = T^{01}dx^2 \wedge dx^3 + T^{02}dx^3 \wedge dx^1 + T^{03}dx^1 \wedge dx^2$.

c) $\dfrac{\partial}{\partial t}\int\limits_V \dfrac{1}{c}T^{\alpha 0}d^3x = -\oint\limits_{\partial V} T^{\alpha\beta}d\sigma_\beta$

称三维向量 $T^{\alpha\beta}$ 为压力张量 (动量流的密度张量).

37.4. 考虑泛函 $S[\Gamma] = \int R\sqrt{|g|}dV$, 其中 $R = g^{ik}R_{ik}, R_{ik} = \dfrac{\partial \Gamma^l_{ik}}{\partial x^l} - \dfrac{\partial \Gamma^l_{il}}{\partial x^k} + \Gamma^l_{ik}\Gamma^m_{lm} - \Gamma^m_{il}\Gamma^l_{km}$, 而度量 g_{ik} 被假定为常值. 证明, 对这个泛函的极值 $P = (\Gamma^k_{ij})$, 克里斯托费尔公式成立, 此公式通过度量张量的分量表达出 Γ^k_{ij}.

37.5. 引力场的能量 - 动量张量是什么 (按 2—4 小节中给出的定义)?

37.6. 设 $F = F_{ik}dx^i \wedge dx^k, i, k = 0, 1, 2, 3$ 是在具度量 (g_{ij}) 的四维时空空间中的电磁场张量. 证明麦克斯韦方程这时有形式 $dF = 0, \delta F = \dfrac{4\pi}{c}j$, 其中 j 为电流,$\delta = *^{-1}d*$.

37.7. 在偶维黎曼 (或伪黎曼) 空间中考虑函数

$$S[g] = \int \Omega,$$

其中形式 Ω 为习题 30.10 中的定义. 证明

$$\frac{\delta S}{\delta g^{ij}} = 0.$$

§38. 拉格朗日的例子

1. 考虑在 $\mathbb{R}_1^4$ 中的复标量场$\varphi(x)$, 其中 $\mathbb{R}_1^4$ 所具有的闵可夫斯基度量为 $g_{ab} = \begin{pmatrix} 1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & -1 \end{pmatrix}$, 我们还给出了一个作用, 其形如

$$S = 常数 \cdot \int \left[\hbar^2 \left\langle \frac{\partial\bar{\varphi}}{\partial x}, \frac{\partial\varphi}{\partial x} \right\rangle - m^2c^2\overline{\varphi(x)}\varphi(x)\right] d^4x = \int \Lambda d^4x, \tag{1}$$

其中的符号 "$-$" 表示复共轭,$\hbar$ 为普朗克常数 (作用的维数),c 为光速. 这里的 $m \geqslant 0$ 被称作质点的质量, 质点由场 φ 描述, $\Lambda = \Lambda\left(\varphi, \bar{\varphi}, \dfrac{\partial\varphi}{\partial x}, \dfrac{\partial\bar{\varphi}}{\partial x}\right)$ 为拉格朗日;$\bar{\varphi}$ 和 φ 被形式地假定为独立变量. 欧拉 - 拉格朗日方程 $\dfrac{\delta S}{\delta\varphi} = 0, \dfrac{\partial S}{\partial\bar{\varphi}} = 0$ 具有 "克莱因 - 戈登方程" 的形式:

$$(\hbar^2\Box + m^2c^2)\varphi = 0, \quad \Box = \frac{\partial^2}{(\partial x^0)^2} - \sum_{\alpha=1}^{3}\frac{\partial^2}{(\partial x^\alpha)^2} \tag{2}$$

($\delta\varphi$ 和 $\delta\bar{\varphi}$ 假定为独立的).

能量 - 动量张量的形式为 (参看 §37)

$$T^{ba} = T^{ab} = \hbar^2 g^{ac} g^{bd} \left(\frac{\partial\bar{\varphi}}{\partial x^c}\frac{\partial\varphi}{\partial x^d} + \frac{\partial\varphi}{\partial x^c}\frac{\partial\bar{\varphi}}{\partial x^d}\right) - g^{ab}\Lambda. \tag{3}$$

能量密度

$$T^{00} = \hbar^2 \sum_{c=0}^{3}\frac{\partial\bar{\varphi}}{\partial x^c}\frac{\partial\varphi}{\partial x^c} + m^2\bar{\varphi}\varphi. \tag{4}$$

作用 (1) 对于变换群

$$\varphi \to e^{i\alpha}\varphi, \quad \bar{\varphi} \to e^{-i\alpha}\bar{\varphi}\ (\alpha = 常数) \tag{5}$$

不变. 这个群生成了它的保有电流:

$$J^a = ig^{ab}\left(\bar{\varphi}\frac{\partial\varphi}{\partial x^b} - \varphi\frac{\partial\bar{\varphi}}{\partial x^b}\right). \tag{6}$$

习题 38.1. 由欧拉 – 拉格朗日方程推导出等式

$$\frac{\partial J^a}{\partial x^a} = 0.$$

称 $\int\limits_{t=\text{常数}} J^0 d^3x = Q$ 为场 φ 的 "电荷".

按 $p \to p + \frac{e}{c}A$, 其中 $p = i\hbar\frac{\partial}{\partial x}$, 产生了电磁场;

$$i\frac{\partial}{\partial x^a} \to i\frac{\partial}{\partial x^a} + \frac{e}{c\hbar}A_a(x), \tag{7}$$

其中 $A_a(x)$ 为电磁场的位势向量,e 为电量,c 和 $\hbar$ 为万有常数. 总拉格朗日为 (设 $\hbar = 1$)

$$\Lambda(\varphi, \bar{\varphi}, A) = \left\langle \frac{\partial\bar{\varphi}}{\partial x^a} + \frac{ie}{c}A_a\bar{\varphi}, \frac{\partial\varphi}{\partial x^a} - \frac{ie}{c}A_a\varphi \right\rangle - m^2c^2\bar{\varphi}\varphi - \frac{1}{16\pi c}F_{ab}F^{ab}. \tag{8}$$

习题 38.2. 验证这个作用在 "度规变换"($\hbar = c = 1$)

$$\varphi \to e^{i\alpha}\varphi, \quad \bar{\varphi} = e^{-i\alpha}\bar{\varphi}, \quad A_a \to A_a + \frac{\partial\alpha}{\partial x^a}, \quad \alpha = \alpha(x), \quad S \to S \tag{9}$$

下的不变性.

实标量场像所说的那样, 是 "中性的"($\bar{\varphi} \equiv \varphi$); 这里的

$$\Lambda = \frac{1}{2}\left\langle \frac{\partial\varphi}{\partial x}, \frac{\partial\varphi}{\partial x} \right\rangle - m^2c^2\varphi^2, \quad (\Box + m^2c^2)\varphi = 0, \tag{10}$$

而向量 (6) 化为零,$J^a = 0$. 不可能出现电磁场, 这是因为欧拉 – 拉格朗日方程的解不会是实的.

自由方程 (2) 或 (10) 的形如常数 $\cdot e^{i(k,x)}$ 的解具有性质

$$\langle k, k\rangle = \frac{m^2c^2}{\hbar^2}. \tag{11}$$

可以认为, 这样的解代表了质量为 m 的自由质点, 其动量为 $p = \hbar k$, 即与向量 k 成比例. 故而动量在具质量的曲面 $\langle p, p\rangle = m^2c^2$ 上.

2. 在 $\mathbb{R}^4_1$ 中质量 m 的复向量场 $\varphi = (\varphi_a)$ 的拉格朗日的形式 (设 $\hbar = c = 1$) 为

$$\Lambda = -g^{bd}g^{ac}\frac{\partial\bar{\varphi}_a}{\partial x^b}\frac{\partial\varphi_c}{\partial x^d} + m^2g^{ab}\bar{\varphi}_a\varphi_b,$$

$$\varphi = (\varphi_0, \varphi_1, \varphi_2; \varphi_3), \quad S = \int \Lambda d^4x, \tag{12}$$

同时加上对场的附加条件 (参看下面的习题 38.4):

$$\frac{\partial\varphi^a}{\partial x^a} = 0, \quad \frac{\partial\bar{\varphi}^a}{\partial x^a} = 0. \tag{13}$$

能量 - 动量张量的形式为

$$T^{ab}=T^{ba}=-g^{ac}g^{bd}g^{kl}\left(\frac{\partial\bar{\varphi}_k}{\partial x^c}\frac{\partial\varphi_l}{\partial x^d}+\frac{\partial\bar{\varphi}_l}{\partial x^d}\frac{\partial\varphi_k}{\partial x^c}\right)-g^{ab}\Lambda. \tag{14}$$

群 $\varphi\to e^{i\alpha}\varphi,\bar{\varphi}\to e^{-i\alpha}\bar{\varphi}$ 生成了所保有的电流

$$J^a=-ig^{ab}g^{cd}\left(\frac{\partial\varphi_c}{\partial x^b}\bar{\varphi}_d-\frac{\partial\bar{\varphi}_c}{\partial x^b}\varphi_d\right). \tag{15}$$

习题

38.3. 验证, 由于欧拉 - 拉格朗日方程

$$\frac{\delta S}{\delta\bar{\varphi}}=0,\quad \frac{\delta S}{\delta\varphi}=0\leftrightarrow(\Box+m^2)\varphi=0, \tag{16}$$

则电流守恒: $\dfrac{\partial J^a}{\partial x^a}=0$.

38.4. 验证, 在对向量场不加附加条件 (13) 时, 能量 $\displaystyle\int T^{00}d^3x$ 一般说来不是一个正的量.

具有平面波类型的解 $e^{i\langle k,x\rangle}$ 有 (当 $\hbar=c=1$) 性质 $\langle k,k\rangle=m^2$ (请验证!). 因此 k 为质量曲面上的动量.

按规则

$$i\frac{\partial}{\partial x^a}\to i\frac{\partial}{\partial x^a}+eA_a\ (\hbar=c=1) \tag{17}$$

重新产生了电磁场. 总拉格朗日 (连同电磁场) 关于下面的度规变换不变:

$$\varphi\to e^{i\alpha(x)}\varphi,\quad \bar{\varphi}=e^{-i\alpha(x)}\bar{\varphi},\quad A_a\to A_a+\frac{\partial\alpha(x)}{\partial x^a}. \tag{18}$$

一个特殊情形是我们以前已经遇到过的零质量 $m=0$ 的向量场 (例如, 电磁场). 这时, 如果场 φ 为实的, 其拉格朗日等价于电磁场的拉格朗日

$$\Lambda=\text{常数}\times F_{ab}F^{ab},\quad \text{其中}F_{ab}=\frac{\partial\varphi_b}{\partial x^a}-\frac{\partial\varphi_a}{\partial x^b}, \tag{19}$$

它关于度规群

$$\varphi_a\to\varphi_a-\frac{\partial\alpha(x)}{\partial x^a},S\to S \tag{20}$$

是不变的.

稍后 (见 §40) 还将考虑 "旋量" 场这个重要的类, 其中场 φ 的值在旋量空间中.

§39. 广义相对论的最简单概念

1. 我们首先回忆爱因斯坦的狭义相对论 (CTO; 有意思的事是在这个理论的创造中, 除了众所周知的物理学家爱因斯坦和洛伦兹外还有当时最杰出的几何学家庞加莱和闵可夫斯基的参与). 根据 CTO, 在某个瞬间, 在曲面的一个点上所发生的"事件"相关联一个四维时空空间 $\mathbb{R}^4_1$ 中的一个点, 其度量为闵可夫斯基度量 $dl^2 = (dx^0)^2 - \sum_{\alpha=1}^{3}(dx^\alpha)^2$, x^α 为伪欧氏坐标, $\alpha = 0,1,2,3$; 其中 $x^0 = ct, t$ 为时间,c 为在真空中的光速 ($c \approx 299\ 793$ km/s). 类时的, 光的 (或迷向的) 和类空的向量 ξ 分别由 $\langle\xi,\xi\rangle > 0, \langle\xi,\xi\rangle = 0$ 或 $\langle\xi,\xi\rangle < 0$ 定义; 而曲线 $\gamma(\tau)$ 为类时的, 光的, 或类空的, 按定义是说速度向量 $v = \dfrac{d\gamma}{d\tau}$ 在每点分别为 $\langle v,v\rangle > 0, \langle v,v\rangle = 0$ 或 $\langle v,v\rangle < 0$. 质点 $(m > 0)$ 的世界线是类时的, 无质量的质点 $(m = 0)$ 的世界线为光的. 在阐述 CTO 时可以利用两个自由质点的拉格朗日中的一个

$$S^{(1)} = \int\limits_{\gamma(\tau)} L^{(1)}_{\mathrm{CB}} d\tau = \alpha \int\limits_{\gamma(\tau)} \langle v,v\rangle d\tau$$

或

$$S^{(2)} = \int\limits_{\gamma(\tau)} L^{(2)}_{\mathrm{CB}} d\tau = \beta \int\limits_{\gamma(\tau)} \sqrt{\langle v,v\rangle} d\tau.$$

这里 $\gamma(\tau)$ 为在闵可夫斯基空间 $\mathbb{R}^4_1$ 中质点的世界线,$v = \dfrac{d\gamma(\tau)}{d\tau}$ 为 4 元速度, α 和 β 为常数, 它们的值 $\alpha = \dfrac{1}{2}mc, \beta = -mc$(参看 §32). 通常在物理文献中对于作用 $S^{(2)} = \int L^{(2)}_{\mathrm{CB}} d\tau$ 使用 $L^{(2)}_{\mathrm{CB}} = \beta\sqrt{\langle v,v\rangle}$, 与世界线 $\gamma(\tau)$ 的四维长度成比例. 这个泛函不依赖于参数 τ. 故可选取 $\tau = \dfrac{x^0}{c} = t$ (世界时). 最后我们得到在三维体系中的拉格朗日, 它适于与经典力学进行对照

$$L^{(2)}_{\mathrm{CB}} = \beta\sqrt{\langle v,v\rangle} = \beta c\sqrt{1 - \frac{|w|^2}{c^2}}, \tag{1}$$

其中 $v = \dfrac{d\gamma}{dt}$ 或 $v^a = \dfrac{dx^a}{d(x^0/c)} = c\dfrac{dx^a}{dx^0}$,

$$v^0 = c, \quad v^a = \frac{dx^a}{dt} = w^a, \quad a = 1,2,3$$

(w^a 为 $\mathbb{R}^3$ 中的所谓三维速度向量 $w = \left(\dfrac{dx^a}{dt}\right)$ 的分量). 如果 $|w| \ll c$ (非相对论的

情形), 则

$$L_{\mathrm{CB}}^{(2)}=\beta\sqrt{\langle v,v\rangle}=\beta c\sqrt{1-\frac{|w|^2}{c^2}}=\beta c\left(1-\frac{w^2}{2c^2}+O\left(\frac{w^4}{c^4}\right)\right). \tag{2}$$

令 $\beta=-mc$, 那么当 $|w|/c\to 0$ 时我们得到在经典力学中自由质点的拉格朗日, 准确到 $O(w^4/c^4)$ 项.

对于自由质点的能量和动量有形式

$$E=w\frac{\partial L_{\mathrm{CB}}^{(2)}}{\partial w}-L_{\mathrm{CB}}^{(2)}=\frac{mc^2}{\sqrt{1-\dfrac{w^2}{c^2}}},\quad p_a=\frac{\partial L_{\mathrm{CB}}^{(2)}}{\partial w^a}=-\frac{mw^a}{\sqrt{1-\dfrac{w^2}{c^2}}}. \tag{3}$$

令 $E/c=p_0$, 我们得到 4 元向量 $p=(p^a)$, 其中 $p^a=g^{ab}p_b, g^{ab}$ 为闵可夫斯基度量. 直接验证质量为 m 的自由质点的 4 元动量所满足的方程

$$p^ap^bg_{ab}=\langle p,p\rangle=(p_0)^2-\sum_{a=1}^{3}(p_a)^2=m^2c^2. \tag{4}$$

方程 (4) 给出在具度量 g_{ab} 的动量 (p^a) 的线性 (余切) 空间中的具质量的曲面. 这就是罗巴切夫斯基空间 (参看 §10), 其上的体积元为 $d\sigma=\dfrac{d^3p}{p^0}$.

在运用拉格朗日 $L_{\mathrm{CB}}^{(1)}=\alpha\langle v,v\rangle$ 时, 我们从 §31 的结果得到: 形式的量 $v\dfrac{\partial L_{\mathrm{CB}}^{(1)}}{\partial v}-L_{\mathrm{CB}}^{(1)}=L_{\mathrm{CB}}^{(1)}=\alpha\langle v,v\rangle$ 沿极值曲线守恒 (因为欧拉 – 拉格朗日方程). 因此 $\langle v,v\rangle=$ 常数; 从而在曲线 $\gamma(\tau)$ 上参数 τ 应该是自然的: $dl=$ 常数 $\times d\tau$. 设 $d\tau=\dfrac{dl}{c}$ (固有时间). 我们定义能量 – 动量的 4 元张量 $p_a=\dfrac{\partial L_{\mathrm{CB}}^{(1)}}{\partial v^a}$, 其中 $v^a=\dfrac{dx^a}{d\tau}$, τ 为自然参数, $d\tau=\dfrac{dl}{c}$. 选取 $\alpha=m/2$, 我们得到对拉格朗日 $L_{\mathrm{CB}}^{(2)}$(上面的) 得到的同一个 4 元动量 $(p_0,p_a)=p$ 的值. 特别, 由此得到 (p^a) 确实是在 $\mathbb{R}_1^4$ 中变换下的 4 元向量. 量 $\langle v,v\rangle$ 沿轨线为常值; 当 $d\tau=\dfrac{dl}{c}$, 我们有 $\langle v,v\rangle=c^2$. 对参数 $\tau=\dfrac{1}{c}$, 通常称量 v 为不变 4 元速度 (v^a), 以和 3 元速度 $\mathbf{w}=(w^1,w^2,w^3)$ 加以区别. 对应关系为

$$w^a=c\frac{v^a}{v^0},\quad \langle v,v\rangle=(v^0)^2-\sum_{a=1}^{3}(v^a)^2=c^2. \tag{5}$$

如果在空间 $\mathbb{R}_1^4$ 有具位势向量 $A_a(x)$ 的电磁场, 则在外部场中质点的拉格朗日为

$$L^{(1)}=L_{\mathrm{CB}}^{(1)}+\frac{e}{c}A_a\frac{dx^a}{d\tau},\quad \tau=\frac{l}{c}, \tag{6}$$

或

$$L^{(2)} = L^{(2)}_{\mathrm{CB}} + \frac{e}{c}A_a\frac{dx^a}{dt} + eA_0, \quad t = \frac{x^0}{c}. \tag{7}$$

因此, 外部场 $A_a(x)$ 的出现等价于 4 元动量的位移 (参看 §33)

$$p_a \to p_a + \frac{e}{c}A_a(x). \tag{8}$$

在哈密顿体系中自由质点的哈密顿有形式 $H = E(p) = c\sqrt{p^2+m^2c^2} = cp_0$, 其中 $p^2 = \sum_{\alpha=1}^{3}(p_\alpha)^2$. 在存在场 $A_a = (A_0, A_\alpha)$ 时, 我们有

$$E(p) \to H(x,p) = c\sqrt{\sum_\alpha (p_\alpha + \frac{e}{c}A_\alpha)^2 + m^2c^2} + eA_0(x) \tag{9}$$

(参看 §33). 场强张量为 $F_{ab} = \frac{\partial A_b}{\partial x^a} - \frac{\partial A_a}{\partial x^b}$, 其中 $F_{0\alpha} = E_\alpha$ (电场),$F_{\alpha\beta} = H_{\alpha\beta}$ (磁场). 这个场的作用为

$$S[A] = -\frac{1}{16\pi c}\int F_{ab}F^{ab}d^4x. \tag{10}$$

这时它们应该满足方程

1) $d(F_{ab}dx^a \wedge dx^b) = \left(\frac{\partial F_{ab}}{\partial x^c} - \frac{\partial F_{ac}}{\partial x^b} + \frac{\partial F_{bc}}{\partial x^a}\right)dx^a \wedge dx^b \wedge dx^c = 0$. (第一对麦克斯韦方程).

2) 在无质点时我们有 $\frac{\partial F^{ab}}{\partial x^b} = 0$ (第二对麦克斯韦方程).

如果有一组质点, 各具电荷 $e_1, \cdots, e_N$, 质量为 $m_1, \cdots, m_N$, 它们在 $\mathbb{R}^4_1$ 中的世界线为 $\gamma_1, \cdots, \gamma_N$, 场为 $A_a(x)$, 于是质点系和场的总作用为 (两种形式)

$$S^{(1)} = \sum_{i=1}^{N}\alpha_i\int_{\gamma_i}\langle v_i, v_i\rangle d\tau + \sum_{i=1}^{N}\int_{\gamma_i}\frac{e_i}{c}(A_a(x)v_i^a)d\tau - \frac{1}{16\pi c}\int F_{ab}F^{ab}d^4x; \tag{11}$$

$$S^{(2)} = -\sum_{i=1}^{N}m_ic^2\int_{\gamma_i}\sqrt{1-\frac{w_i^2}{c^2}}dt + \sum_{i=1}^{N}\int_{\gamma_i}\left[\frac{e_i}{c}A_\alpha(x)w_i^\alpha dt + e_iA_0(x)dt\right] - \frac{1}{16\pi c}\int F_{ab}F^{ab}d^4x, \tag{12}$$

其中 $\gamma_i(\tau) = (x_i^a(\tau))$ 为第 i 个质点坐标, $w_i = (w_i^\alpha)$ 和 $v_i = (v_i^a)$ 为第 i 个质点的速度. 由在外部场 (11) 中质点的拉格朗日的形式 $L^{(1)} = \alpha\langle v, v\rangle + \frac{e}{c}A_av^a$, 得到在外部场中沿极值线的参数与自然参数 (或固有时间)

$$v\frac{\partial L^{(1)}}{\partial v} - L^{(1)} = v\frac{\partial L^{(1)}_{\mathrm{CB}}}{\partial v} - L^{(1)}_{\mathrm{CB}} = \alpha\langle v, v\rangle, \quad \frac{d}{d\tau}\langle v, v\rangle = 0.$$

2. 已经讨论过的情形中, 不可能在 CTO 中包含有引力. 更准确地说: 可以纯形式地引进引力的 "位势向量 A_a^G, 其中电荷 e_i 被质量 m_i 替代 (质点的质量是它的电荷引力); 位势向量 $A_a^G(x)$ 是这样取的, 在我们需定义了引力场极值的特殊坐标系中, 它应该有形式 $A^G=(\varphi,0,0,0)$, 其中 φ 为通常的引力位势, 它同时满足达朗贝尔方程和拉普拉斯方程. 这时质点在外部场 A^G 中的运动可由 (11) 的拉格朗日得到, 其中 A 换作 A^G, e 换作 m. 但是 (庞加莱已经指出过) 对于水星的近日点的扰动是由对牛顿定律的不确切之处的修定得到的, 它们在数值上与观测所得到的不相合 (虽说这些观测量的阶是正确的). 因此应该寻求另一种方法以将引力场与相对论联系起来. 爱因斯坦的广义相对论(OTO) 的基本假定是这样的: 引力场只不过是度量 g_{ab}, 其符号差在四维时空空间 $M^4(x^0,x^1,x^2,x^3)$ 中为 $(+---)$; 这时, 一般说来, 度量 g_{ab} 有非零曲率 (曲率的大小对引力场非平凡性的程度作了特征描述). 在外引力场中的一个试验质点只不过是在 "具度量 g_{ab} 的空间中的自由质点", 它沿类时测地线 $\gamma(\tau)=(x^a(\tau))$ 运动, 而此测地线则是由下面的拉格朗日 (这里的 $m\neq 0$) 决定

$$L_{\mathrm{CB}}^{(1)}=m\langle v,v\rangle=m\frac{dx^a}{d\tau}\frac{dx^b}{d\tau}g_{ab}. \tag{13}$$

(如果 $m=0$, 则质点将按光的测地线运动, 度量为 g_{ab}.) 沿曲线 $\gamma(\tau)$ 的固有时间为 $\frac{l}{c}=\tau, \frac{dl}{c}=d\tau, |v|=$ 常数.

具位势向量 $A_a(x)$ 的电磁场的存在就如上面那样, 只是闵可夫斯基度量被换成了度量 g_{ab}

$$S=\int\limits_{\gamma(\tau)}L^{(1)}d\tau=\int\limits_{\gamma(\tau)}L_{\mathrm{CB}}^{(1)}d\tau+\frac{e}{c}\int\limits_{\gamma(\tau)}A_a(x)dx^a. \tag{14}$$

这个场的作用有形式

$$S_f=-\frac{1}{16\pi c}\int F_{ab}F^{ab}\sqrt{-g}d^4x,\quad F_{ab}=\frac{\partial A_b}{\partial x^a}-\frac{\partial A_a}{\partial x^b}, \tag{15}$$

其中 $d\sigma=\sqrt{-g}d^4x$ 为体积元, $g=\det(g_{ab}), F^{ab}=g^{ac}g^{bd}F_{cd}$ 为在度量 g_{ab} 下提升张量 F_{ab} 指标的结果. 在度量 g_{ab} 下的麦克斯韦方程 (不存在电荷) $\frac{\delta S_f}{\delta A_a}=0$ 为

$$d(F_{ab}dx^a\wedge dx^b)=0\ (\text{第一对方程}),$$

$$\nabla_b F^{ab}=0\ (\text{第二对方程}). \tag{16}$$

暂时不讨论这个引力场 g_{ab} 的方程, 我们来指出爱因斯坦假设的一个简单的推论. 我们考虑一个充分弱的引力场, 即度量 g_{ab} (我们暂时不对词 "弱" 作精确定义), 以及在此场中一个 "慢的", 带质量的质点, 它沿着测地线 $x(\tau)$ 运动,

$$\ddot{x}^a+\Gamma_{bc}^a\dot{x}^b\dot{x}^c=0,\quad a,b,c=0,1,2,3, \tag{17}$$

其中 Γ^{a}_{bc} 为克氏符号 (由度量 g_{ab} 产生的联络的系数, 见 §29). 时间在这里取为沿测地线的固有时间即自然参数 $\tau = l/c$. 设 $t = x^0/c$ 如同在 CTO 中那样. 我们假定"弱" 度量被表示为 $1/c$ 的形式级数, 这时设 $1/c$ 为小的参数

$$g_{ab} = g^{(0)}_{ab} + c^{-2}g^{(2)}_{ab} + c^{-3}g^{(3)}_{ab} + \cdots = g^{(0)}_{ab} + O\left(\frac{1}{c^2}\right) \tag{18}$$

(假定了 $g^{(1)}_{ab} = 0$, 而 $g^{(0)}_{ab}$ 为闵可夫斯基度量). 我们假定质点是慢的, 即 $\left|\frac{dx^a}{dt}\right| \ll c$ (按 "慢" 的定义, 量 $\frac{1}{c}\frac{dx^a}{dt}$ 为 $O\left(\frac{1}{c}\right)$ 阶). 对于固有时间 τ 或者说是自然参数有

$$d\tau = \frac{dl}{c} = \sqrt{\frac{1}{c^2}\frac{dx^a}{dt}\frac{dx^b}{dt}g_{ab}}\quad dt\ \left(t = \frac{x_0}{c}\right) \tag{19}$$

或者

$$d\tau = \sqrt{1 + O\left(\frac{1}{c^2}\right)}dt = \left[1 + O\left(\frac{1}{c^2}\right)\right]dt. \tag{20}$$

对克里斯托费尔符号我们有 (见 §29)

$$\Gamma^{a}_{bc} = \frac{1}{2}g^{ad}\left(\frac{\partial g_{bd}}{\partial x^c} + \frac{\partial g_{cd}}{\partial x^b} - \frac{\partial g_{bc}}{\partial x^d}\right). \tag{21}$$

因为 $x^0 = ct, t$ 为有限, 于是由 (18) 在 Γ^{a}_{bc} 的表达式中的微分 $\frac{\partial g_{ab}}{\partial x^0}$ 的量级为 $O\left(\frac{1}{c^3}\right)$ (闵可夫斯基度量 $g^{(0)}_{ab}$ 为常值)

微分 $\frac{\partial g_{ab}}{\partial x^\alpha}$ 的阶为 $O\left(\frac{1}{c^3}\right)$, $\alpha = 1, 2, 3$, 这是因为 x^α 被认为是有限的. 在方程 (17) 中, 因 (20) 可以以同样的精度以 dt 替换 $d\tau$; 在这些方程 $(\alpha = 1, 2, 3)$ 中有克里斯托费尔符号的项的最高阶为 $\Gamma^{\alpha}_{00}\dot{x}^0\dot{x}^0 = \Gamma^{\alpha}_{00}c^2 + O\left(\frac{1}{c}\right)$, 这是因为当 $\alpha = 1, 2, 3$ 时量 $\dot{x}^\alpha$ 具有阶 $O(1)$. 因 (18), 对于 Γ^{α}_{00} 有

$$\Gamma^{\alpha}_{00} = \frac{1}{2}g^{\alpha\alpha}\left(-\frac{\partial g_{00}}{\partial x^\alpha}\right) + O\left(\frac{1}{c^3}\right),$$

或者

$$\Gamma^{\alpha}_{00} = \frac{1}{2}\frac{\partial g_{00}}{\partial x^\alpha} + O\left(\frac{1}{c^3}\right). \tag{22}$$

那么方程 (17) 便有了形式

$$\frac{d^2x^\alpha}{dt^2} = -\frac{1}{2}\frac{\partial g_{00}}{\partial x^\alpha}c^2 + O\left(\frac{1}{c}\right). \tag{23}$$

对在弱场中的慢质点, 方程 (23) 应该与在具势能 $\varphi(x)$ 的经典引力场中质点运动相符合, 其精度为 $O\left(\frac{1}{c}\right)$ 量级. 于是 g_{00} 的系数应该是

$$g_{00} = 1 + \frac{2\varphi(x)}{c^2} + O\left(\frac{1}{c^3}\right), \tag{24}$$

其中 φ 为牛顿引力位势. 这时, 方程 (23) 得到了牛顿方程的样子

$$\frac{d^2x^\alpha}{dt^2} = -\frac{\partial\varphi}{\partial x^\alpha} + O\left(\frac{1}{c}\right). \tag{25}$$

因此成立

命题 39.1. 如果爱因斯坦假设成立, 则度量 g_{00} 应以下面的关系与引力位置相关 (在弱场中)

$$g_{00} = 1 + \frac{2\varphi(x)}{c^2} + O\left(\frac{1}{c^3}\right).$$

特别, 固有时间 $d\tau = \frac{dl}{c}$ 与世界时间 dt 不同; 对于固定不动的质点, 如其世界线为 $x^\alpha =$ 常数 $(\alpha = 1,2,3)$, 则我们得到

$$cd\tau = \sqrt{g_{00}\left(\frac{dx^0}{dt}\right)^2}dt = c\sqrt{g_{00}}dt, \quad d\tau \approx dt\left(1 + \frac{\varphi(x)}{c^2}\right). \tag{26}$$

因为总有 $\varphi \leqslant 0$, 我们可推出: 在弱引力场中, 两个事件之间的固有时间 (对固定不动质点) 小于它们之间的世界时间

$$\text{如果 } \varphi < 0, \text{则 } \tau < t.$$

(我们有两个事件 $(x^0 = x_1^0, x^\alpha = 0)$ 和 $(x^0 = x_2^0, x^\alpha = 0)$ 沿世界线 $x^\alpha = 0, \alpha = 1,2,3$ 之间的时间区间)

3. 现在考虑引力场 (g_{ab}), 即在四维空间中区域上的度量有符号差 $(+---)$, 其中没有其他的场, 也没有质点. 让我们给出假定: 引力场应是所谓 "万有不变的", 即这个引力场的理论应在所有坐标系中有唯一的形式, 并且它通过度量 g_{ab} 的曲率张量 R^a_{bcd} 表出. 我们将不对此问题进行详细讨论, 而叙述通过里奇曲率 $R_{bc} = R^a_{bac}$ 给出的在真空中引力场的爱因斯坦方程

$$R_{ab} - \frac{1}{2}Rg_{ab} = 0 \tag{27}$$

(或者因为 $\frac{1}{2}R = \frac{1}{2}R^a_a$, 有 $R_{ab} = 0$). 爱因斯坦算子 $R_{ab} - \frac{1}{2}Rg_{ab}$ 对任意度量 g_{ab} 具有下面的重要性质

$$\nabla_a(R^a_b - \frac{1}{2}R\delta^a_b) \equiv 0 \tag{28}$$

(比安基恒等式的推论, 参看习题 30.8). 关于度量 g_{ab} 的方程 (27) 具有下列性质:

1) 这个方程的阶数为 2.

2) 对于弱场, $g_{00}=1+\frac{2\varphi(x)}{c^2}+O\left(\frac{1}{c^3}\right)$, 此方程化为对势能 φ 的泊松方程:

$$\Delta\varphi=\sum_{\alpha=1}^{3}\frac{\partial^2\varphi(x)}{(\partial x^\alpha)^2}=0 \tag{29}$$

(将在下面阐述).

3) 这个方程可以被写成欧拉–拉格朗日方程. 其作用及拉格朗日有形式 (由希尔伯特证明)

$$S=\int R\sqrt{-g}d^4x,\quad g=\det(g_{ab}),\quad R=R^a_a. \tag{30}$$

变差 δS 和变分导数 $\frac{\partial S}{\partial g^{ab}}$ 已在 §37 中计算, 从而有形式

$$\delta S=\int\left(R_{ab}-\frac{1}{2}Rg_{ab}\right)\delta g^{ab}\sqrt{-g}d^4x,$$

$$\frac{1}{\sqrt{-g}}\frac{\partial S}{\partial g^{ab}}=R_{ab}-\frac{1}{2}Rg_{ab}. \tag{31}$$

习题 39.1. 证明形如 $R_{ab}-\gamma Rg_{ab}$ 的算子是变分导数当且仅当 $\gamma=\frac{1}{2}$; 在拉格朗日情形中应满足等式 (28).

注 形如下面的方程

$$R_{ab}-\frac{1}{2}Rg_{ab}=\lambda g_{ab} \tag{32}$$

满足 1)—3) 的性质. 迄今没有根据认为 $\lambda\neq 0$(称 λ 为 "宇宙常数"; 姑且认为 $\lambda=0$; 因依宇宙学估值有 $\lambda<10^{-56}\text{cm}^{-2}$).

定常的球面式的对称解

$$\varphi=\frac{\text{常数}}{r},\quad r^2=\sum_{\alpha=1}^{3}(x^\alpha)^2 \tag{33}$$

是最简单的泊松方程 $\Delta\varphi=0$(在所建立的质量场之外) 的非平凡解, 在那里常数等于物体质量的总和, 而这些物体构成了场 φ(物体为球面对称).

$$\varphi=-G\frac{M}{r}, \tag{34}$$

其中 $G=6.67\cdot10^{-8}\text{cm}^3/(\text{g}\cdot\text{s})$ 为牛顿引力常数. 我们求在 OTO 中类比的牛顿位势 $-G\frac{M}{r}$. 我们考虑球面对称的度量 g_{ab}, 它不依赖于时间 $t=\frac{x^0}{c}$. 寻找形如

$$dl^2 = g_{00}c^2dt^2 + g_{11}dr^2 - r^2d\Omega^2 \tag{35}$$

的度量, 其中 $g_{00} = e^{\nu} > 0, g_{11} = -e^{\lambda} < 0, d\Omega^2 = d\theta^2 + \sin^2\theta d\varphi^2$ 为单位球面在球坐标 θ, φ 下的长度元. 这时, 由定常性和球面对称性得到 $g_{00} = g_{00}(r), g_{11} = g_{11}(r)$. 对于系数 Γ^a_{bc}, 按照克里斯托费尔公式 (参看 §29)($x^0 = ct, x^1 = r, x^2 = \theta, x^3 = \varphi$), 我们有 ($a' = \frac{\partial a}{\partial r}$)

$$\Gamma^1_{11} = \frac{\lambda'}{2}, \quad \Gamma^0_{10} = \frac{\nu'}{2}, \quad \Gamma^2_{33} = -\sin\ \theta\cos\ \theta, \quad \Gamma^1_{33} = -r\sin^2\theta e^{-\lambda},$$

$$\Gamma^1_{22} = -re^{-\lambda}, \quad \Gamma^1_{00} = \frac{\nu'}{2}e^{\nu-\lambda}, \quad \Gamma^2_{12} = \Gamma^3_{13} = \frac{1}{r}, \quad \Gamma^3_{23} = \cot\ \theta.$$

建立爱因斯坦方程$R_{ab} = 0$, 其中

$$R_{ab} = \frac{\partial \Gamma^c_{ab}}{\partial x^c} - \frac{\partial \Gamma^c_{ac}}{\partial x^b} + \Gamma^c_{ab}\Gamma^d_{cd} - \Gamma^c_{ad}\Gamma^d_{bc},$$

那么我们得到

$$\lambda = 0, \quad e^{-\lambda}\left(\frac{\nu'}{r} + \frac{1}{r^2}\right) - \frac{1}{r^2} = 0, \quad e^{-\lambda}\left(\frac{\nu'}{r} - \frac{1}{r^2}\right) + \frac{1}{r^2} = 0.$$

我们有积分 $\lambda + \nu = f(t)$. 经替换 $t' = \psi(t)$ 可使函数 f 为零. 解的形式为

$$g_{00} = 1 - \frac{a}{r}, \quad g_{11} = -\frac{1}{1-(a/r)},$$

其中 a 为某个常数.

下面是施瓦氏度量

$$dl^2 = \left(1 - \frac{a}{r}\right)c^2dt^2 - \frac{1}{1-(a/r)}dr^2 - r^2d\Omega^2. \tag{36}$$

当 $r \to \infty$ 时度量的扰动变弱, 从而我们有

$$g_{ab} \approx g^{(0)}_{ab} + \frac{1}{c^2}g^{(2)}_{ab} + O\left(\frac{1}{c^3}\right), \tag{37}$$

其中

$$g^{(2)}_{00} = -\frac{ac^2}{r}. \tag{38}$$

于是, 比较 (24),(34), 我们得到 $a = \frac{2MG}{c^2}$, 其中 M 为构成场的物体的质量,G 和 c 为常数. 称对物体质量 M 的量 $a = \frac{2GM}{c^2}$ 为引力的施瓦氏半径(对地球的质量 $a = 0.44$ cm, 对太阳的质量 $a = 3$ km). 如果物体如此稠密, 使其维数的级为 (或小

于)a, 于是由公式 (36) 显然当 $r \to a$ 时有某些奇异性发生. 关于这些奇异性在第二卷 §31 中有更详细的阐述. 目前可以相信的只是该公式 (36) 在区域 $r > a$ 是正确的.

我们早先曾指出, 系数 g_{00} 由位势 φ 定义 (准确到 $O\left(\frac{1}{c^3}\right)$). 对于把慢的类时测地线和在场中质点的经典的牛顿方程进行比较, 这已经足够了 (见前面). 在完全知道了场的度量时, 我们也可以研究在大的 r 时快速质点的运动进行修正, 特别是无质量的质点. 在度量 (36) 中光的测地线方程有形式

$$\frac{d^2x^a}{d\eta^2} + \Gamma^a_{bc}\frac{dx^a}{d\eta}\frac{dx^b}{d\eta} = 0, \quad a = 0, 1, 2, 3, \tag{39}$$

其中 $g_{ab}\dfrac{dx^a}{d\eta}\dfrac{dx^b}{d\eta} = 0$. 不解出方程, 我们在这里将光的测地线公式化为球对称的施瓦氏度量 (测地线位于坐标 r, φ 的平面中). 光测地线的方程为

$$\varphi = \int \frac{dr}{r^2\sqrt{\dfrac{1}{\rho^2} - \dfrac{1}{r^2}\left(1 - \dfrac{a}{r}\right)}}, \quad \rho = \text{常数}. \tag{40}$$

当 $a \to 0$ 时, 作为极限, 我们得到直线 $r\cos\varphi = \rho$. 对于小的 a 可以由 $\varphi(r)$ 的公式得到光线离直线的修正值.

4. 按照 OTO 的观点, 假设了在所有场和质点 (除了引力场以外的所有场!) 与度量 g_{ab}(引力量) 之间存在着相互作用, 它是通过所谓的能量 – 动量张量 T_{ab} 以下面的方式产生的: 爱因斯坦的总方程为

$$R_{ab} - \frac{1}{2}Rg_{ab} = \text{常数} \cdot T_{ab} \tag{41}$$

或者

$$R^b_a - \frac{1}{2}R\delta^b_a = \text{常数} \cdot T^b_a. \tag{42}$$

这里的常数被假定为万有的 (普遍适用的). 对于其更精确的值, 应考虑形如 (36) 的弱度量, 和尘云的能量 – 动量张量, 其中的压力等于零, 物质的速度等于零. 这时这个张量的形状为

$$T_{ab} = \begin{pmatrix} \rho c^2 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & 0 & \\ 0 & & & \end{pmatrix}, \tag{43}$$

其中 ρ 为质量的密度. 我们知道 $g_{00} = 1 + \dfrac{2\varphi}{c^2} + O\left(\dfrac{1}{c^3}\right)$, 其中 φ 为引力位势, 它满足泊松方程

$$\Delta\varphi = 4\pi G\rho \tag{44}$$

我们注意到, 对张量 (43) 有 $T_0^0=\rho c^2, T_a^a=\rho c^2$. 在模去 $O\left(\frac{1}{c^2}\right)$ 后的非平凡量 Γ_{bc}^a, 对于计算 R_0^0 是必须知道的, 它们为

$$\Gamma_{00}^{\alpha}=\frac{1}{c^2}\frac{\partial\varphi}{\partial x^{\alpha}}+O\left(\frac{1}{c^3}\right),$$

$$R_0^0=\frac{1}{c^2}\left(\sum_{\alpha=1}^{3}\frac{\partial^2\varphi}{(\partial x^{\alpha})^2}\right)+O\left(\frac{1}{c^3}\right)=\frac{1}{c^2}\Delta\varphi+O\left(\frac{1}{c^3}\right).$$

爱因斯坦方程有

$$R_0^0-\frac{1}{2}R\delta_0^0=常数\cdot T_0^0,$$

因而

$$R_0^0=常数\cdot(T_0^0-\frac{1}{2}T_a^a)=常数\cdot\frac{\rho c^2}{2}. \tag{45}$$

对于能量 – 动量张量 (43) 我们有 $R_0^0=\frac{\Delta\varphi}{c^2}=常数\cdot\frac{\rho c^2}{2}$. 要求爱因斯坦方程 (45) 当能量 – 动量张量具形式 (43) 时转换为泊松方程 (44), 则可以得出结论:

$$R_{ab}-\frac{1}{2}Rg_{ab}=\frac{8\pi G}{c^4}T_{ab}, \tag{46}$$

这时我们已假定知道了万有常数 (即 $\frac{8\pi G}{c^4}$). 由关系 (46) 得到了等式

$$\nabla_b T_a^b=0, \tag{47}$$

它代替了 OTO 中的守恒定律.

我们将给出一些能量 – 动量张量: 我们曾遇到过, 而且是极其重要的.

1) 电磁场的能量 – 动量张量(参看前面)

$$T^{ab}=\frac{1}{4\pi c}(-F^{ac}F_c^b+\frac{1}{4}g^{ab}F_{ad}F^{cd}). \tag{48}$$

2) 连续介质的能量 – 动量张量 (流体动力学的能量 – 动量张量)

$$T_{ab}=(p+\varepsilon)u_a u_b-pg_{ab}, \tag{49}$$

其中 p 和 ε 分别为介质在所在的 "相伴" 坐标系中 (在给定点上) 的压力和能量密度 (或说 $u_0=1, u_\alpha=0, \alpha=1,2,3$). 这里的 $u=\frac{v}{c}$ 为介质的 4 元速度向量, 满足 $\langle u,u\rangle=u_a u_b g^{ab}=1$. 在所说的相伴参照系中,$T_{ab}$ 的形式为

$$T_{ab}=\begin{pmatrix}\varepsilon & & & 0\\ & p & & \\ & & p & \\ 0 & & & p\end{pmatrix}. \tag{50}$$

为了完全确定爱因斯坦方程, 我们必须要知道介质的状态方程, 即 p 和 ε 之间的关联. 对于尘云物质有 $p=0, \varepsilon=\rho c^2$. 对于所谓的 "超相对论" 状态方程有 $p=\varepsilon/3$ 或 $T_a^a \equiv 0$.

习题 39.2. 证明电磁场的能量 - 动量张量的形式为

$$\frac{1}{2}T_{ab}=\frac{1}{\sqrt{-g}}\frac{\delta S}{\delta g^{ab}},$$

其中

$$S=-\frac{1}{16\pi c}\int F_{ab}F_{cd}g^{ac}g^{bd}\sqrt{-g}d^4x, \quad g=\det(g_{ab}).$$

注 1. 在阐述 OTO 时通常按下面方式定义物质的对称的能量 - 动量张量:

$$\frac{\delta S_m}{\delta g^{ab}}=\frac{1}{2}\sqrt{-g}T_{ab}, \tag{51}$$

其中对物质场的作用被假定为由既依赖于物质场也依赖于引力场度量的泛函形式的显式表达.

例如, 对于标量和向量场的拉格朗日 (见 §38), 度量显式地含在拉格朗日中梯度的内积形式里. 对于旋量场 (参看 §40) 事情远为复杂. 在 OTO 中引进旋量将在 §41 中讨论. 在这一章中所指出的连续介质的能量 - 动量张量显式地包含了度量. 引力场和物质的总的作用具有的形式为

$$S=\int(R\sqrt{-g}d^4x+S_m). \tag{52}$$

例如, 如果引力场 F_{ab} 是 "物质的", 则

$$S=\int(R+\text{常数}\cdot F_{ab}F^{ab})\sqrt{-g}d^4x. \tag{53}$$

完全的方程组 (麦克斯韦和爱因斯坦的) 的形式为

$$\frac{\delta S}{\delta A_a}=0, \quad \frac{\delta S}{\delta g^{ab}}=0. \tag{54}$$

2. 曾经多次讨论过这个引力场的能量密度有什么准确意义的问题 (以及将所有的能量 - 动量张量当作万有共变的量的问题). 由上面说过的, 在取 (51) 为在一般物理上唯一正确的定义时, 我们看到在这个关系中, 引力场有别于所有其他类型的物质场, 在这些物质场中能量 - 动量张量是由相对于所给出的度量而定义的 (当然是认定了爱因斯坦基本假定的正确性, 即认为引力场就是度量; 这个结论现在已为一系的实验成功验证). 所有这些讨论的结果大体是说, 除了爱因斯坦方程的第一部分外, 不存在任何万有的共变能量 - 动量张量. 有一个使 "整体引力能量" (非稠密的!) 有意义的重要情形: 在闵可夫斯基度量的背景下考虑 "局部的" 纯引力团. 由与经典的非相对论的重力相比较得到结论说, 度量的分量 (准确地说, 它们与闵可夫斯基度

量的离差) 应该充分正则地以阶为 r^{-1} 的速度 (不会更快!) 当 $r \to \infty$ 时按空间的方向减小. 这时总的 "质量场" 通过其对空间的无限性的逼近来定义; 逼近的规则是通过对照在经典的引力中的势能逼近得出相应的物体质量的定义. 作为对三维度量和它的时间导数的泛函, 这个量原来是个哈密顿系统, 因而可以被看成为局部引力团的物理能量. 这个量为正的性质直到最近才严格地从物理和几何上得到证明. 这个 "引力能量" 对于任意 (内部的) 坐标变换不变, 但是对这个定义的合理性和唯一性应该以足够快的速度消退在空间的无穷远处. 值得留意的是, 虽然局部 "引力团" 已经按闵可夫斯基度量决定, 但在这种情形中, 按定义把它看作是在闵可夫斯基度量的背景下的一个新的对象还是恰当的, 而且它的能量也按照同一个一般物理的规则被确定. 现在还不知道有一个这样局部的解存在 (无度量的奇异性, 没有其他类型的物质), 甚至它们的存在性显然也不能有效地从现有的文献中的数学定理中提取出来; 这样的 "局部非线性引力波" 暂时还看不出来, 而且也难于迅速发现; 但是, 这个问题自身有很大的形式上的兴趣, 并且在现代文献中有着激烈的讨论. 还有一种表示在任意背景附近的小振动理论的情形, 这可以看成是一种场论, 它的能量是按照背景的度量定义的 (但是在构造这种理论时必须小心, 并且要引进关于消除坐标任意性的相容规则).

这些问题在已知书本的范围内还没有被详细地被考虑过.

§40. 群 $SO(3)$ 和 $O(3,1)$ 的旋量表示. 狄拉克方程和它的性质

1. 矩阵代数的自同构

我们考虑在 n 维空间 $\mathbb{C}^n$ 中的整个的矩阵代数 $M(n,\mathbb{C})$. 所考虑的群的旋量表示的构造所根据的是矩阵代数的下面性质. 我们记得, 代数的自同构 $h: M(n,\mathbb{C}) \to M(n,\mathbb{C})$ 是指它到自身的同构; 称形如 $h(x) = gxg^{-1}$ 的自同构为代数 $M(n,\mathbb{C})$ 的内自同构, 其中 g 为群 $GL(n,\mathbb{C})$ 中的矩阵.

引理 40.1. 结合代数 $M(n,\mathbb{C})$ 的任意自同构都是内自同构.

证明 我们记得代数 $M(n,\mathbb{C})$ 的元 P 被称为射影 (算) 子(幂等元) 是说 $P^2 = P$, 称射影算子 P 和 Q 为正交是说 $PQ = QP = 0$; 显然, 正交的射影算子的像相交为零. 如果 $P(\mathbb{C}^n)$ 为一维空间, 则称射影算子为一维算子. 特别, 由公式

$$(P_i)^k_l = \begin{cases} 1, \text{如果 } k = l = i, \\ 0, \text{如果} k \neq i \text{或} l \neq i \end{cases} \tag{1}$$

定义的矩阵是两两正交的一维射影算子. 我们有

$$P_i^2 = P_i, \quad P_i P_j = 0 \text{ 当 } i \neq j, \quad P_1 + \cdots + P_n = 1. \tag{2}$$

现在考虑自同构 $h: M(n,\mathbb{C}) \to M(n,\mathbb{C})$, 并令 $h(P_i) = P'_i$. 显然, 因为 $h(1) = 1$, 关系式 (2) 在作用 h 下化为关系式

$$(P'_i)^2 = P'_i, \quad P'_iP'_j = 0 \text{当} i \neq j, \quad P'_1 + \cdots + P'_n = 1. \tag{3}$$

因此,P'_i 为射影算子; 这些算子是非平凡的, 而且两两正交, 从而 $P'_1(\mathbb{C}^n) + \cdots + P'_n(\mathbb{C}^n) = \mathbb{C}^n$. 故而所有这些射影算子是一维的. 令 $P'_i(\mathbb{C}^n) = \mathbb{C}'_i$.

以 t_{ij} 代表代数 $M(n,\mathbb{C})$ 中的元,$1 \leqslant i \leqslant n, 1 \leqslant j \leqslant n$, 其中 $(t_{ij})^k_l = 1$, 当 $i = k, j = l$; $(t_{ij})^k_l = 0$, 对其余 k, l 值; 换句话说,$t_{ij}(e_j) = e_i, t_{ij}(e_r) = 0$ 当 $r \neq j$, 其中 $e_1, \cdots, e_n$ 为空间 $\mathbb{C}^n$ 中的标准基. 显然,

$$t_{ii} = P_i, \quad t_{ij}t_{rs} = 0 \text{ 当} j \neq r, \quad t_{ij}t_{js} = t_{is} \tag{4}$$

(最后面的那个公式不是对 j 求和). 令 $h(t_{ij}) = t'_{ij}$, 并将 h 用于关系式 (3) 便得到关系

$$t'_{ii} = P'_i, \quad t'_{ij}t'_{rs} = 0 \text{ 当} j \neq r, \quad t'_{ij}t'_{js} = t'_{is}. \tag{5}$$

因为当 $k \neq i$ 时 $P'_k t'_{ij} = 0$, 像 $t'_{ij}(\mathbb{C}^n)$ 为一维从而与 $\mathbb{C}'_i$ 相合. 进而, 由于

$$t'_{ij}(\mathbb{C}^n) = t'_{ij}P'_j(\mathbb{C}^n) = t'_{ij}(\mathbb{C}'_j),$$

则 t'_{ij} 同构于 $\mathbb{C}'_j$ 到 $\mathbb{C}'_i$ 的映射.

我们固定一个任意的非零向量 $e'_1 \in \mathbb{C}'_1$ 并令 $e'_i = t'_{i1}(e'_1)$. 向量 $e'_1, \cdots, e'_n$ 不同于零并分别在 $\mathbb{C}'_1, \cdots, \mathbb{C}'_n$ 上; 因此它们构成了空间 $\mathbb{C}^n$ 中的基. 我们以公式

$$g(e_i) = e'_i, \tag{6}$$

定义变换 $g: \mathbb{C}^n \to \mathbb{C}^n$, 并表明了对任意 $x \in M(n,\mathbb{C})$ 有

$$h(x) = gxg^{-1}. \tag{7}$$

我们首先注意到

$$t'_{ij}(e'_j) = t'_{ij}t'_{j1}(e'_1) = t'_{i1}(e'_1) = e'_i,$$

$$t'_{ij}(e'_r) = t'_{ij}t'_{r1}(e'_1) = 0, \text{当} r \neq j.$$

由此显见有 $t'_{ij}(e'_r) = gt_{ij}g^{-1}(e'_r)$, 其中 r 为任意, 即

$$h(t_{ij}) = gt_{ij}g^{-1}$$

对任意 i, j 成立. 但是因为任意矩阵 x 可以表示为形如 t_{ij} 的矩阵的线性组合, 于是 $h(x) = gxg^{-1}$.

引理 40.1 得证. □

某些特殊的矩阵代数 $M(2,\mathbb{C}), M(4,\mathbb{C})$ 可用一些特殊形状的生成元实现, 它们对我们来说是重要的, 称它们分别为泡利矩阵和狄拉克矩阵.

2. 群 $SO(3)$ 的旋量表示

在代数 $M(2,\mathbb{C})$ 中我们选取下列的生成元组:

$$1,\quad \sigma_x=\sigma_1=\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad \sigma_y=\sigma_2=\begin{pmatrix}0&-i\\i&0\end{pmatrix},\quad \sigma_z=\sigma_3=\begin{pmatrix}1&0\\0&-1\end{pmatrix}. \tag{8}$$

称矩阵 $\sigma_j(j=1,2,3)$ 为泡利矩阵(参看 §14); 连同矩阵 1 它们给出了代数 $M(2,\mathbb{C})$ 中的加法基, 这是因为 $M(2,\mathbb{C})$ 作为线性 (四维) 空间时, 它们在其中线性无关. 它们相关联的关系式为

① $\sigma_q\sigma_l-\sigma_l\sigma_q=2i\sigma_k$, 其中 (q,l,k) 为偶置换;

② $\sigma_q\sigma_l+\sigma_l\sigma_q=2\delta_{ql}$.

我们将它们写为

① $[\sigma_q,\sigma_l]=2i\varepsilon_{qlk}\sigma_k$,

② $\{\sigma_q,\sigma_l\}=2\delta_{ql}$.

关系 ① 简单地表明矩阵 $\frac{i}{2}\sigma_j$ 实现了群 $SO(3)$ 的李代数表示 (或由定理 24.3, 为群 $SU(2)$ 的李代数).

我们特别留意一下关系 ② 的一个性质. 我们把三维欧氏空间 $\mathbb{R}^3(x^1,x^2,x^3)$ 实现为无迹 (迹为 0) 的 2×2 矩阵的空间, 其欧氏基为 σ_j

$$(x^1,x^2,x^3)\longmapsto x^\alpha\sigma_\alpha.$$

设 $\Lambda\in O(3)$ 为正交变换

$$\Lambda:\mathbb{R}^3\to\mathbb{R}^3,\quad x^{\alpha'}=\lambda^\alpha_\beta x^\beta. \tag{9}$$

令

$$\sigma'_q=\lambda^\alpha_q\sigma_\alpha. \tag{10}$$

原来, 关系式 (2) 在形如 (10) 的正交变换下保持不变

$$\{\sigma'_q,\sigma'_l\}=2\delta_{ql}.$$

(这可由矩阵 Λ 的正交性定义以显见的方式推导出来.) 另一方面, 对关系式 ① 我们有

$$\begin{aligned}[\sigma'_\alpha,\sigma'_\beta]&=[\lambda^\gamma_\alpha\sigma_\gamma,\lambda^\delta_\beta\sigma_\delta]=\lambda^\gamma_\alpha\lambda^\delta_\beta[\sigma_\gamma,\sigma_\delta]\\&=2i(\lambda^\gamma_\alpha\lambda^\delta_\beta\varepsilon^t_{\gamma\delta}\sigma_t),\end{aligned}$$

其中

$$\varepsilon^t_{\gamma\delta}=\varepsilon_{\gamma\delta t}=\begin{cases}1, & \text{如果 } (\gamma,\delta,t) \text{ 为偶置换},\\ -1, & \text{如果 } (\gamma,\delta,t) \text{ 为非偶置换},\\ 0, & \text{如果 } \gamma,\delta,t \text{ 中有重复的}.\end{cases}$$

最后有

$$[\sigma'_\alpha, \sigma'_\beta] = 2i\varepsilon_{\alpha\beta\gamma}\sigma'_\gamma, \text{ 如果 } A \in SO(3).$$

因此, 关系式 ① 对形如 (10) 的变换下保持不变, 其中在 $\mathbb{R}^3$ 中的正交矩阵 A 的行列式为 $+1$. 关系式 ① 和 ② 完全由矩阵代数 $M(2,\mathbb{C})$ 确定. 事实上, 由于这些关系式, 所有的乘积 $\sigma_i\sigma_j$ 被 $1, \sigma_1, \sigma_2, \sigma_3$ 线性表出. 于是得到

定理 40.1. 对 $SO(3)$ 中 A 的变换 (10) 给出了一个自同构 $h(A) : M(2,\mathbb{C}) \to M(2,\mathbb{C})$. 因此, 由引理 40.1 存在变换 $g = g(A) : \mathbb{C}^2 \to \mathbb{C}^2$ 使得 $h(x) = gxg^{-1}$, 其中 $x \in M(2,\mathbb{C})$ 为任意的矩阵.

定义 40.1. 称关系 $A \mapsto g(A)$ 为群 $SO(3)$ 到群 $GL(2,\mathbb{C})$ 中的*旋量表示*.

这个表示是多值的: 矩阵 $g(A)$ 确定到差一个非零因子 $\lambda \in \mathbb{C}$. 如果我们要求 $g(A) \in SL(2,\mathbb{C})$, 则此不确定性可以消减. 于是得到了一个双值表示, 它两个值相互差一个 -1 因子. 故而旋量表示是由群 $SO(3)$ 到群 $SL(2,\mathbb{C})/\pm 1$ 的同态.

习题 40.1. 证明后面那个双值表示的像位于 $SU(2)$ 中, 这个双值表示与投影 $SU(2) \to SU(2)/\pm 1$ 的复合可作为同构实现

$$SO(3) \to SU(2)/\pm 1.$$

验证绕以向量 $n = (n_x, n_y, n_z), n_x^2 + n_y^2 + n_z^2 = 1$ 为方向的轴的旋转对应于变换

$$g(n, \varphi) = \exp\left\{-i\frac{\varphi}{2}(n_x\sigma_x + n_y\sigma_y + n_z\sigma_z)\right\}.$$

3. 洛伦兹群的旋量表示

转而讨论代数 $M(4,\mathbb{C})$. 在这个代数中选取生成元 $1, \gamma^0, \gamma^1, \gamma^2, \gamma^3$, 它们满足关系

$$\gamma^a\gamma^b + \gamma^b\gamma^a = 2g^{ab} \cdot 1, \tag{11}$$

其中 g^{ab} 为闵可夫斯基度量, 为此只要取

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \gamma^2 = \begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix}, \gamma^1 = \begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix}, \gamma^3 = \begin{pmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{pmatrix} \tag{12}$$

(这是个 4×4 矩阵, 可写成 2×2 矩阵的 2×2 块矩阵.)

引理 40.2. 所有下面的 4×4 矩阵 $1, \gamma^a, \gamma^a\gamma^b (a < b), \gamma^a\gamma^b\gamma^c (a < b < c)$ 和 $\gamma^5 = \gamma^0\gamma^1\gamma^2\gamma^3$ 为线性无关; 域 $\mathbb{C}$ 上由 $\gamma^0, \gamma^1, \gamma^2, \gamma^3$ 及关系 (11) 生成的代数同构于矩阵代数 $M(4,\mathbb{C})$.

证明 由关系式 (11) 知, 元素 γ^a 的任意乘积都是这个引理所指出的那些元素的线性组合. 另外, 这些生成元的个数为 16, 正好是 $M(4,\mathbb{C})$ 的维数. 因此, 如果验证了引理中由满足 (12) 的矩阵 $\gamma^0, \gamma^1, \gamma^2, \gamma^3$ 的乘积组成元的线性无关性, 则此引理中的所有论断都将得以证明. 我们假定读者已经写出了所有这 16 个矩阵并直接验证了它们的线性无关性 (这里并无复杂的计算). □

现在转而讨论群 $O(3,1)$ 的旋量表示的构造, 为此我们要运用引理 40.1. 考虑闵可夫斯基空间 $\mathbb{R}^4(x^0,x^1,x^2,x^3)$ 以及矩阵 $\Lambda\in O(3,1),\Lambda=(\lambda_b^a)$. 定义矩阵

$$\gamma'^a=\lambda_b^a\gamma^b. \tag{13}$$

由关系式 (11), 并由矩阵 Λ 保持闵可夫斯基度量 g_{ab} 不变的事实得到 γ'^a 满足同一个关系式

$$\{\gamma'^a,\gamma'^b\}=2g^{ab}\cdot 1.$$

由引理 40.2, 这表明映射 $h=h(\Lambda):M(4,\mathbb{C})\to M(4,\mathbb{C})$ 为整个线性代数 $M(4,\mathbb{C})$ 的自同构, 其中这个映射按公式 $1\mapsto 1,\gamma^a\mapsto\gamma'^a$ 进行. 因此存在这样的 $g=g(\Lambda)\in GL(4,\mathbb{C})$, 使得 $h(x)=gxg^{-1},x$ 为 $M(4,\mathbb{C})$ 中所有的元素. 称对应 $\Lambda\mapsto g(\Lambda)$ 为群 $O(3,1)$ 在群 $GL(4,\mathbb{C})$ 中的*旋量表示*. 这个表示是多值的: 矩阵 g 和 λg, 其中 λ 为非零复数, 它们都对应于同一个 $\Lambda\in O(3,1)$. 过渡到群 $SL(4,\mathbb{C})$, 我们得到了双值表示

$$\Lambda\mapsto\pm g(\Lambda)\in SL(4,\mathbb{C}).$$

定义 40.2. 称旋量表示所作用的四维空间 $\mathbb{C}^4$ 为 (4 元分量) 旋量空间. 称空间中的元 $\psi\in\mathbb{C}^4$ 为旋量 (写成列向量).

群 $O(3,1)$ 的旋量表示在子群 $SO(3)$ 上的限制分解为两个同构于上面所描述的不可约旋量表示的和.

注 选取基 $\bar\gamma^0=\gamma^0,\bar\gamma^a=-i\gamma^a$, 我们得到关系式

$$\{\bar\gamma^a,\bar\gamma^b\}=2\delta^{ab}\cdot 1. \tag{14}$$

利用它们可以构建群 $SO(4)$ 的旋量表示, 它类似于群 $O(3,1)$ 的旋量表示.

习题

40.2. 对于 $SO(3)\subset O(3,1)$ 中绕 $\mathbb{R}^3$ 中单位向量 $n=(n_1,n_2,n_3)$ 以角 φ 的旋转, 其旋量表示为

$$g(\varphi,n)=\exp\left\{-i\frac{\varphi}{2}(n_1\Sigma_1+n_2\Sigma_2+n_3\Sigma_3)\right\},$$

其中

$$\Sigma_j=\begin{pmatrix}\sigma_j&0\\0&\sigma_j\end{pmatrix}.$$

40.3. 对于在平面 (x^0,n) 上以虚角 $i\varphi$ 的旋转, 其中 n 为三维单位向量 (基本洛伦兹变换), 旋量表示为

$$g(\varphi,n)=\exp\left\{-\frac{\varphi}{2}(n_1\alpha_1+n_2\alpha_2+n_3\alpha_3)\right\},其中\alpha_j=\begin{pmatrix}0&\sigma_j\\\sigma_j&0\end{pmatrix}.$$

40.4. 证明空间映射 $P(x^0,x)=(x^0,-x)$ 的表示为

$$g(P)=\eta_P\gamma^0, \text{其中} \eta_P=\pm i \text{ 或 } \eta_P=\pm 1,$$

而 x 为三维向量.

40.5. 证明时间映射算子 $T(x^0,x)=(-x^0,x)$ 的表示为

$$g(T)=\eta_T\gamma^0\gamma^1\gamma^3, \text{其中} |\eta_T|=1.$$

由构造 (见公式 (12)) 知, 空间 $\mathbb{C}^4$ 分解为和: $\mathbb{C}^4=\mathbb{C}^2\oplus\mathbb{C}^2$,

$$\psi=\begin{pmatrix}\varphi\\ \chi\end{pmatrix},\quad \varphi=\mathbb{C}^2,\quad \chi\in\mathbb{C}^2. \tag{15}$$

在旋量的空间中,$\psi=\begin{pmatrix}\varphi\\ \chi\end{pmatrix}$ 可以化为新的基

$$\eta=\frac{\varphi+\chi}{\sqrt{2}},\quad \xi=\frac{\varphi-\chi}{\sqrt{2}}. \tag{16}$$

由给出旋量表示的公式 (习题 1—3) 知道,"半旋量"η 和 ξ 在特殊洛伦兹 $SO(3,1)\subset O(3,1)$ 中的变换进行变换时是独立的, 并且在空间映射(请验证.)

$$\begin{aligned} &g(P):\eta\mapsto\xi,\xi\mapsto\eta,\\ &P(x^0,x)=(x^0,-x)\end{aligned} \tag{17}$$

相互转换.

定义 40.3. 群 $SO(3,1)$ 在 η 和 ξ 上的作用被称做特殊洛伦兹群 $SO(3,1)$ 的半旋表示;记 η 为g_+, ξ 为 g_-,(对半旋表示的其他记号将在 §41.3 给出).

这些表示 (各自单独地) 不能拓展到整个洛伦兹群 $O(3,1)$.

我们注意, 在基 $\psi=\begin{pmatrix}\eta\\ \xi\end{pmatrix}$ 下矩阵 γ^0 有形式 (请验证!)

$$\gamma^0=\begin{pmatrix}0&1\\1&0\end{pmatrix}. \tag{18}$$

群 $O(3,1)$ 的旋量表示不是酉表示. 但是, 可以构造一个不定内积, 它对于旋量表示不变: 为此只要令 $\langle\psi,\psi\rangle=\psi^*\gamma^0\psi$. 这里的 $\psi^*=(\psi_1^*,\psi_2^*,\psi_3^*,\psi_4^*)$ 为共轭于列向量 ψ 的行向量. 在基 $\psi=\begin{pmatrix}\eta\\ \xi\end{pmatrix}$ 下, 形式 $\langle\psi,\psi\rangle=\bar\psi\psi$ 有形式 $\bar\psi\psi=\psi^*\gamma^0\psi=\xi^*\eta+\eta^*\xi$(请验证!).

习题 40.6. 验证 $\psi^*\gamma^0\psi$ 在群 $O(3,1)$ 下不变.

定义 40.4. (行) 旋量 $\bar\psi=\psi^*\gamma^0$ 被称做 ψ 的狄拉克共轭(我们注意到, 形式 $\bar\psi\psi$ 对旋量表示不变.)

习题 40.7. 证明量 $\bar{\psi}\gamma^a\psi$ 按向量那样变化 (关于 $O(3,1)$), 量 $\bar{\psi}\gamma^a\gamma^b\psi$ 为 2 阶张量, $\bar{\psi}\gamma^a\gamma^b\gamma^c\psi$ 为 3 阶张量,$\bar{\psi}\gamma^5\psi = \bar{\psi}\gamma^0\gamma^1\gamma^2\gamma^3\psi$ 为 4 阶张量 (伪标量)

40.8. 证明群 $SO(3,1)$ 在 $GL(2,\mathbb{C})$ 的半旋量表示 g_+ 和 g_- 不具有任何不变的非零内积.

提示 在半旋量表示里不存在不变性是因为群 $SL(2,\mathbb{C})$ 中的标准表示中的一个为同构, 而其余的为共轭表示.

群 $SL(2,\mathbb{C})$ 的李代数与洛伦兹群之间的同构为同构 (24.72) 在 $n=2$ 时的特殊情形, 其中球面 S^2 的共形变换的李代数是作为群 $SO(3,1)$ 的李代数实现的.

4. 狄拉克方程

γ 矩阵在关系 (11) 下构成的代数自然产生下面的问题: 设给出了克莱因—戈登算子 $\Box + m^2$, 其中 $\Box = g^{ab}\dfrac{\partial}{\partial x^a}\dfrac{\partial}{\partial x^b}$, 可将其展开为 1 阶算子的乘积吗?

$$-(\Box + m^2) = \left(i\gamma^a \frac{\partial}{\partial x^a} + m\right)\left(i\gamma^b \frac{\partial}{\partial x^b} + (-m)\right). \tag{19}$$

习题 40.9. 证明展开式 (19) 等价于关系

$$\{\gamma^a, \gamma^b\} = 2g^{ab}\cdot 1 \tag{20}$$

(即如果取 γ 矩阵代数中元素作为系数, 则展式 (19) 是可以得到的).

定义 40.5 称在旋量 ψ 上的方程

$$\left(i\gamma^b \frac{\partial}{\partial x^b} - m\right)\psi = 0 \tag{21}$$

为*狄拉克方程*.

习题 40.10. 验证共轭方程的形式为

$$i\frac{\partial\bar{\psi}}{\partial x^a}\gamma^a + m\bar{\psi} = 0, \tag{22}$$

其中 $\bar{\psi} = \psi^*\gamma^0$.

狄拉克方程由作用

$$S = \int \left[\frac{i}{2}\left(\bar{\psi}\gamma^a \frac{\partial\psi}{\partial x^a} - \frac{\partial\bar{\psi}}{\partial x^a}\gamma^a\psi\right) - m\bar{\psi}\psi\right] d^4x \tag{23}$$

得出, 其中 ψ 和 ψ 被假定是无关的. 能量—动量张量和电流有形式

$$T^{ab} = \frac{i}{2}g^{ac}\left(\bar{\psi}\gamma^b \frac{\partial\psi}{\partial x^c} - \frac{\partial\bar{\psi}}{\partial x^c}\gamma^b\psi\right) = T^{ba}, \tag{24}$$

$$J^a = \bar{\psi}\gamma^a\psi, \ a = 0,1,2,3. \tag{25}$$

量 $J^0 = \bar{\psi}\gamma^0\psi = \psi^*(\gamma^0)^2\psi = \psi^*\psi$ 为电荷密度. 电荷的形式为

$$Q = \int \psi^*\psi d^3x = \int J^0 d^3x. \tag{26}$$

显然,$Q > 0$. 相反地, 能量密度 T^{00} 则不是正定的; 这引起了一系列的困难; 我们将不在这里讨论它.

类型为常数 $\cdot e^{i\langle k,x\rangle}$ 的解在质量平面 $\langle k,k\rangle = m^2(\hbar = c = 1)$ 具有波向量. 如果 $k^0 > 0$(即动量位于质量曲面的上面部分), 它们则恒同于质点. 这个条件的必要性出于对质点能量的正定性要求.

在基 $\{\eta,\xi\}$ 中狄拉克方程取得的形式为

$$i\frac{\partial\eta}{\partial t} = \langle\sigma,p\rangle\eta + m\xi, \quad i\frac{\partial\xi}{\partial t} = -\langle\sigma,p\rangle\xi + m\eta, \tag{27}$$

其中 $p = i\dfrac{\partial}{\partial x}$ $(\hbar = c = 1)$ 和 $\langle\sigma,p\rangle = i\left(\sigma_1\dfrac{\partial}{\partial x^1} + \sigma_2\dfrac{\partial}{\partial x^2} + \sigma_3\dfrac{\partial}{\partial x^3}\right)$. 我们看到, 当质量 $m = 0$ 时, 这个方程分成了两个独立的方程 (外尔方程), 它们描述了其运动规律在空间反射时不是不变的质点 (群 $SO(3,1)$ 的半旋表示不能拓展到空间的反射上), 并且它具有质量零 (半旋表示不具有不变的内积, 这个内积在拉格朗日上可以给出质量型的项).

5. 电磁场的狄拉克方程. 电荷的共轭算子

依照标准的规则: $p_a \to p_a + eA_a(\hbar = e = 1)$ 给出了外部电磁场出现的情形. 因此我们应该作通常的拉格朗日中的变换以及方程

$$\frac{\partial}{\partial x^a} \to \frac{\partial}{\partial x^a} - ieA_a(x),$$

其中 e 为电荷 $(\hbar = e = 1)$. 在外部场中狄拉克方程的形式为

$$\left[\tilde{\gamma}^a\left(\frac{\partial}{\partial x^a} - ieA_a(x)\right) + m\right]\psi = 0, \quad \bar{\psi}\left[(\tilde{\gamma}^a)^{\mathrm{T}}\left(\frac{\partial}{\partial x^a} + ieA_a(x)\right) - m\right] = 0, \tag{28}$$

其中引进的是一组埃尔米特矩阵 $\tilde{\gamma}^\alpha : \tilde{\gamma}^0 = \gamma^0, \tilde{\gamma}^\alpha = -i\gamma^\alpha, \alpha = 1,2,3$ (参照第三小节的变换), “T” 表示转置.

考虑 (实) 矩阵 $C = \tilde{\gamma}^2\tilde{\gamma}^0$. 它具有下面的性质 (请验证!):

$$C^{-1}\tilde{\gamma}^a C = -(\tilde{\gamma}^a)^{\mathrm{T}}, \quad a = 0,1,2,3. \tag{29}$$

由 (29) 和狄拉克方程得到

定理 40.2. 如果场 $\psi(x),\psi(x)$ 满足狄拉克方程 (28), 它在场 $A_a(x)$ 中电荷为 e, 于是场 $\psi^c(x) = C\bar{\psi}(x)^{\mathrm{T}}$, 并且 $\bar{\psi}^c(x) = (C^{-1}\psi(x))^{\mathrm{T}}$ 在同一个场 $A_a(x)$ 中满足方程 (28), 但是将 e 替换成 $-e$.

称变换 $\psi \to \psi^c$ 为电荷的共轭变换, 这是因为描述场 ψ 的质点所具有电荷的符号改变了. 由此得到了重要的推论: 旋量表示和狄拉克方程直接描述了两类质点: 具电荷 e 的质点和具电荷 $-e$ 的质点. 狄拉克方程 (28) 的同一个解 $\psi(x)$ 定义了电子波动函数 $\psi(x)$ 和正电子的波动函数 $\psi^c(x)$.

习题 40.11. 验证:a) 电荷共轭算子 $\psi \to \psi^c$ 与特殊洛伦兹群可交换; b) 这个算子与空间反射 $g(P) = \eta_P \tilde{\gamma}^0$ 交换, 条件是 $\eta_P = \pm i$(如果 $\eta_P = \pm 1$, 则不可交换); c) 如果 $\eta_P = \pm 1$, 这个算子与时间反射交换 (参看习题 40.4,40.5).

§41. 具有任意对称性的场的共变微分

1. 度规变换. 度规不变的拉格朗日

由重要的例子开始. 我们考虑复标量场 ψ 和形如

$$L = L\left(\psi, \bar{\psi}, \frac{\partial \psi}{\partial x^\alpha}, \frac{\partial \bar{\psi}}{\partial x^\alpha}\right) \tag{1}$$

的拉格朗日, 其中 "杠" 表示复共轭,ψ 和 $\bar{\psi}$ 被假定为形式上无关的 (见 §38).

设拉格朗日 L 对于形如

$$\psi \to e^{i\varphi}\psi, \quad \bar{\psi} \to e^{-i\varphi}\bar{\psi} \ (\varphi\text{为常数}), \tag{2}$$

$$L\left(e^{i\varphi}\psi, e^{-i\varphi}\bar{\psi}, e^{i\varphi}\frac{\partial \psi}{\partial x^\alpha}, e^{-i\varphi}\frac{\partial \bar{\psi}}{\partial x^\alpha}\right) = L\left(\psi, \bar{\psi}, \frac{\partial \psi}{\partial x^\alpha}, \frac{\partial \bar{\psi}}{\partial x^\alpha}\right). \tag{3}$$

的电荷的变换群下不变: 例如, 在 §38 中所考虑过的形如

$$L = \hbar^2 g^{\alpha\beta}\frac{\partial \bar{\psi}}{\partial x^\alpha}\frac{\partial \psi}{\partial x^\beta} - m^2c^2\bar{\psi}\psi \tag{4}$$

的拉格朗日便具有这样的不变性. 由局部性的原则出发, 我们要求拉格朗日关于更广泛的形如 (2) 的不变性, 就是说, 我们允许 φ 依赖于 x. 这表明群 (2) 独立地作用在每个点上.

原来, 用下面的过程可以得到对于变换

$$\psi \to e^{ie\varphi(x)}\psi, \quad \bar{\psi} \to e^{-ie\varphi(x)}\bar{\psi} \tag{5}$$

下不变的拉格朗日:

1) 引入新的余向量场 A_α, 它在变换 (5) 下以梯度的变换进行变化

$$A_\alpha \to A_\alpha - \frac{\partial \varphi}{\partial x^\alpha} \ (\hbar = c = 1); \tag{6}$$

2) 引入新的拉格朗日 $\widetilde{L}$,

$$\widetilde{L}\left(\psi, \psi, \frac{\partial \psi}{\partial x^\alpha}, \frac{\partial \bar{\psi}}{\partial x^\alpha}, A_\alpha\right) = L(\psi, \bar{\psi}, \nabla_\alpha\psi, \overline{\nabla_\alpha\psi}), \tag{7}$$

其中

$$\nabla_\alpha \psi = \frac{\partial \psi}{\partial x^\alpha} + ieA_\alpha \psi; \tag{8}$$

表达式 (7) 是个标量, 这是因为 $\nabla_\alpha \psi$ 为余向量;

3) 不变理论的总拉格朗日应有形式

$$L(\psi, \bar{\psi}, \nabla\psi, \overline{\nabla\psi}) + L_1\left(A, \frac{\partial A}{\partial x}\right),$$

其中 L_1 为梯度不变的.L_1 这一项的形式将在 §42 中讨论.

定理 41.1. 拉格朗日 (7) 关于 (局部) 变换 (5),(6) 不变.

证明 作为"共变导数"的变换的量, 我们有

$$\begin{aligned}\nabla_\alpha \psi \ \to & \frac{\partial}{\partial x^\alpha}(e^{ie\varphi(x)}\psi) + ie\left(A_\alpha - \frac{\partial \varphi}{\partial x^\alpha}\right)\psi \\ &= e^{ie\varphi(x)}\left[\frac{\partial \psi}{\partial x^\alpha} + ieA_\alpha \psi\right] = e^{ie\varphi(x)}\nabla_\alpha \psi.\end{aligned}$$

因此在变换 (5),(6) 下新的格拉朗日没有变化

$$\begin{aligned}&L(e^{ie\varphi}\psi, e^{-ie\varphi}\bar{\psi}, e^{ie\varphi}\nabla_\alpha\psi, e^{-ie\varphi}\overline{\nabla_\alpha\psi}) \\ &= L(\psi, \bar{\psi}, \nabla_\alpha\psi, \overline{\nabla_\alpha\psi}),\end{aligned}$$

这是因为原来的拉格朗日的不变性. 定理得证. □

我们现在考虑一般情形, 这时 $\psi(x) = (\psi^1(x), \cdots, \psi^N(x))$ 为空间 $\mathbb{R}^n$ 中的场, 其值在 N 维实的向量空间. 设已给出拉格朗日 $L\left(\psi, \dfrac{\partial \psi}{\partial x^\alpha}\right)$, 它对某 N 阶矩阵群 G 不变

$$L\left(g\psi, g\frac{\partial \psi}{\partial x^\alpha}\right) = L\left(\psi, \frac{\partial \psi}{\partial x^\alpha}\right), g \in G. \tag{9}$$

场 $\psi(x)$ 的共变导数由公式

$$\nabla_\alpha \psi = \frac{\partial \psi}{\partial x^\alpha} + A_\alpha(x)\psi \tag{10}$$

给出, 其中 $A_\alpha(x)$ $(\alpha = 1, \cdots, n)$ 为在空间 $\mathbb{R}^n$ 中一组矩阵值函数, 其值取在群 G 的李代数中. 要求量 $A_\alpha(x)$ 在 x 的变换下为余向量那样变化

$$A_\alpha(x) = A_\beta(y)\frac{\partial y^\beta}{\partial x^\alpha}, \quad y = y(x). \tag{11}$$

定理 41.2. 在形如

$$\psi(x) \to g(x)\psi(x), \tag{12}$$

$$A_\alpha(x) \to g(x)A_\alpha(x)g^{-1}(x) - \frac{\partial g(x)}{\partial x^\alpha}g^{-1}(x) \tag{13}$$

的变换下, 其共变导数的变化规律为

$$\nabla_\alpha\psi \to g(x)\nabla_\alpha\psi. \tag{14}$$

在所指出的这个变换下, 拉格朗日 $\widetilde{L}\left(\psi, \frac{\partial\psi}{\partial x^\alpha}, A_\alpha\right) = L(\psi, \nabla_\alpha, \psi)$ 不变.

证明　由公式 (10),(12),(13) 我们得到

$$\begin{aligned}\nabla_\alpha\psi(x) \to\ & \frac{\partial}{\partial x^\alpha}(g(x)\psi) + \left[gA_\alpha g^{-1} - \frac{\partial g(x)}{\partial x^\alpha}g^{-1}\right]g\psi \\ &= g(x)\left[\frac{\partial\psi}{\partial x^\alpha} + A_\alpha\psi\right] = g(x)\nabla_\alpha\psi.\end{aligned}$$

于是等式 (14) 得证. 现在我们注意, 由于 (11), 向量值函数 ψ 的共变导数构成了个余向量

$$\nabla_\alpha\psi(x) = \nabla_\beta\psi(y)\frac{\partial y^\beta}{\partial x^\alpha}, \quad y = y(x),$$

它就像梯度 $\frac{\partial\varphi}{\partial x^\alpha}$ 一样. 故而由等式 (14) 和不变性条件 (9) 得到新拉格朗日的不变性

$$L(g\psi, g\nabla_\alpha\psi) = L(\psi, \nabla_\alpha\psi).$$

得证. □

称所引进的场 A_α 为规范(或补偿)场(或联络), 而称形如 (13) 的变换为度规变换(或规范变换).这些变换构成了规范群.

对于形如 (12) 的局部变换群, 拉格朗日的不变性条件确定了场 ψ 与规范场之间的相互作用, 但并没有这个规范场的拉格朗日. 我们将在 §42 中回到此问题.

注　在群 G 的李代数 $\mathfrak{g}$ 中取值的场是向量值场的一个特殊情形. 设 $B(x)$ 为这样的场. 规范场 (联络)A_α 由公式

$$\nabla_\alpha B = \frac{\partial B}{\partial x^\alpha} + [A_\alpha, B] \tag{15}$$

定义了场 $B(x)$ 的共变导数, 其中的 $[A_\alpha, B] = A_\alpha B - BA_\alpha$ 为李代数中的换位子.

习题

41.1.　验证, 在形如

$$B(x) \to g(x)B(x)g^{-1}(x),$$

$$A_\alpha(x) \to g(x)A_\alpha(x)g^{-1}(x) - \frac{\partial g(x)}{\partial x^\alpha}g^{-1}$$

的度规变换下, 共变导数 $\nabla_\alpha B$ 按规则

$$\nabla_\alpha B \to g(x)(\nabla_\alpha B)g^{-1}(x)$$

变化.

41.2. 验证, 无穷小度规变换可以有形如

$$\psi(x) \to \psi(x) + B(x)\psi(x) + o(B),$$

$$A_\mu(x) \to A_\mu(x) + \nabla_\mu B(x) + o(B)$$

的表示, 其中 $B(x)$ 为在李代数中取值的场.

2. 曲率形式

我们计算两个共变导数的换位子

$$\begin{aligned}[\nabla_\mu, \nabla_\nu]\psi &= \nabla_\mu\nabla_\nu\psi - \nabla_\nu\nabla_\mu\psi \\ &= \frac{\partial}{\partial x^\mu}\left(\frac{\partial\psi}{\partial x^\nu} + A_\nu\psi\right) + A_\mu\left(\frac{\partial\psi}{\partial x^\nu} + A_\nu\psi\right) - \\ &\quad \frac{\partial}{\partial x^\nu}\left(\frac{\partial\psi}{\partial x^\mu} + A_\mu\psi\right) - A_\nu\left(\frac{\partial\psi}{\partial x^\mu} + A_\mu\psi\right) \\ &= \left(\frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu} + [A_\mu, A_\nu]\right)\psi.\end{aligned}$$

引进记号

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu} + [A_\mu, A_\nu]. \tag{16}$$

结论 算子 ∇_μ, ∇_ν 的换位子是乘以矩阵 $F_{\mu\nu}$ 的算子.

定理 41.3. 在度规变换 (13) 下, 量 $F_{\mu\nu}$ 按下面的方式变化

$$F_{\mu\nu} \to gF_{\mu\nu}g^{-1}. \tag{17}$$

量 $F_{\mu\nu}$ 构成了反称二阶张量 (取值于李代数 $\mathfrak{g}$), 并且在坐标变换下按下面规则变化.

$$F_{\mu\nu}(x) = F_{\alpha\beta}(y)\frac{\partial y^\alpha}{\partial x^\mu}\frac{\partial y^\beta}{\partial x^\nu}, \quad y = y(x). \tag{18}$$

证明 直接由替换公式 (13) 到 $F_{\mu\nu}$ 的表达式 (16) 中得到. □

称取值于群 G 的李代数的形式 $\Omega = \sum\limits_{\mu<\nu} F_{\mu\nu}dx^\mu \wedge dx^\nu$ 为联络 A_μ 的曲率形式.

如果存在函数 $g_0(x)$, 它取值于群 G 使得

$$A_\mu(x) = -\frac{\partial g_0(x)}{\partial x^\mu}g_0^{-1}(x), \tag{19}$$

则称联络 A_μ 为平凡联络.

定理 41.4. 平凡联络的曲率形式为零.

证明 对场 A_μ 运用度规变换 $A_\mu \to \bar{A}_\mu = gA_\mu g^{-1} - \dfrac{\partial g}{\partial x^\mu}g^{-1}, g = g_0^{-1}$, 我们得到 $\widetilde{A}_\mu \equiv 0$. 根据前面的论断, 曲率形式 $F_{\mu\nu}$ 变到了形式 $g_0^{-1}F_{\mu\nu}g_0$, 而它等于零. 论断得证. □

习题 41.3. 1) 证明相反的论断: 如果联络的曲率形式恒等于零, 则联络为 (局部) 平凡.

2) 我们定义形式 Ω 的共变微分为

$$D\Omega = \sum_{\lambda<\mu<\nu} (\nabla_\lambda F_{\mu\nu} + \nabla_\nu F_{\lambda\mu} + \nabla_\mu F_{\nu\lambda}) dx^\lambda \wedge dx^\mu \wedge dx^\nu. \tag{20}$$

证明, 如果 Ω 为某个联络的曲率形式, 则比安基恒等式成立:$D\Omega \equiv 0$.

3. 基本例子

a) $G = U(1) \simeq SO(2)$. 我们已经知道 (见 §38), 引进与复标量场相互作用的电磁场等价于在拉格朗日中把每个导数 $\dfrac{\partial\psi}{\partial x^\alpha}$ 替换成共变导数 $\nabla_\alpha\psi = \dfrac{\partial\psi}{\partial x^\alpha} + ieA_\alpha\psi$. 我们从这例子的讨论开始这一节的. 这里的群 G 为一维交换群 $G = U(1) = \{e^{ie\varphi}\}$. 它的李代数也是一维和可交换的 (即有恒为零的换位子), 并且由纯虚数组成. 其联络就是这个电磁场的位势向量 $ieA_\mu(x)$. 这个联络的曲率形式是电磁场的电压张量 $F_{\mu\nu} = \dfrac{\partial A_\nu}{\partial x^\mu} - \dfrac{\partial A_\mu}{\partial x^\nu}$ (在这里无换位子的项, 也略去了因子 ie). 比安基恒等式为形式 $\Omega = \sum\limits_{\mu<\nu} F_{\mu\nu}dx^\mu \wedge dx^\nu = d(A_\mu dx^\mu)$ 为闭的条件.

另一个例子与非交换规范群有关.

b) 线性联络.$G = GL(n,\mathbb{R})$. 在某区域 U 上的坐标 $x^1, \cdots, x^n$ 给出了向量空间中的基 $\left(\dfrac{\partial}{\partial x^1}, \cdots, \dfrac{\partial}{\partial x^n}\right)$. 故而在区域 U 上切向量场可以看作是向量值函数 $\xi = (\xi^1, \cdots, \xi^n)$(在 $\mathbb{R}^n$ 中取值). 坐标变换 $x^\nu \to x^{\nu'} = x^{\nu'}(x)$ 在区域 U 上由形如

$$\xi^\nu \to \xi^{\nu'} = \frac{\partial x^{\nu'}}{\partial x^\nu}\xi^\nu = (g(x)\xi)^{\nu'} \tag{21}$$

的局部变换定义. 这里的矩阵 $g(x) = \left(\dfrac{\partial x^{\nu'}}{\partial x^\nu}\right)$ 可逆, 即在群 $GL(n,\mathbb{R})$ 中. 逆矩阵的形式为 $g^{-1} = \left(\dfrac{\partial x^\nu}{\partial x^{\nu'}}\right)$. 群 $GL(n,\mathbb{R})$ 的李代数全部由 n 阶矩阵组成. 故而联络 $A_\mu(x)$ 自身的系数也是 n 阶矩阵. 我们以 $(A_\mu)^\nu_\lambda = \Gamma^\nu_{\lambda\mu}$ 表示其矩阵的分量. 向量函

数 (即切向量)ξ 的共变导数为

$$\nabla_\mu \xi = \frac{\partial \xi}{\partial x^\mu} + A_\mu \xi, \text{即 } (\nabla_\mu \xi)^\nu = \frac{\partial \xi^\nu}{\partial x^\mu} + \Gamma^\nu_{\lambda\mu}\xi^\lambda. \tag{22}$$

由坐标变换给出的联络分量 $\Gamma^\nu_{\lambda\mu}$ 的度规变换为 $\left(\text{其中 } g(x) = \left(\frac{\partial x^{\nu'}}{\partial x^\nu}\right)\right)$

$$\Gamma^\nu_{\lambda\mu} \to \Gamma^{\nu'}_{\lambda'\mu} = \frac{\partial x^{\nu'}}{\partial x^\nu}\Gamma^\nu_{\lambda\mu}\frac{\partial x^\lambda}{\partial x^{\lambda'}} + \frac{\partial x^{\nu'}}{\partial x^\nu}\frac{\partial}{\partial x^\mu}\left(\frac{\partial x^\nu}{\partial x^{\lambda'}}\right). \tag{23}$$

因为 $A_\mu(x)$ 为余向量, 则 $A_{\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}}, A_\mu$, 即

$$\Gamma^{\nu'}_{\lambda'\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial x^{\nu'}}{\partial x^\nu}\Gamma^\nu_{\lambda\mu}\frac{\partial x^\lambda}{\partial x^{\lambda'}} + \frac{\partial x^{\nu'}}{\partial x^\nu}\frac{\partial^2 x^\nu}{\partial x^{\lambda'}\partial x^{\mu'}}. \tag{24}$$

我们曾在 §28 中推导出克里斯托费尔符号的变化规律. 我们特别留意到这个联络的曲率的矩阵形式为

$$(F_{\mu\nu})^\kappa_\lambda = R^\kappa_{\lambda,\mu\nu}, \tag{25}$$

其中 $R^\kappa_{\lambda,\mu\nu}$ 为曲率张量 (见 §30).

设在区域 U 中给出了黎曼度量 $g_{\alpha\beta}(x)$. 构造 (局部) 两两正交的切向量的法化基 $\xi_1, \cdots, \xi_n$:

$$\langle \xi_\alpha, \xi_\beta \rangle = \delta_{\alpha\beta}.$$

对于与度量 $g_{\alpha\beta}$ 相容的对称联络的形如 $\nabla_{\xi_\alpha}\xi_\beta = \sum \widehat{\Gamma}_{\beta\alpha\gamma}\xi_\gamma$ 的共变导数曾在 §29.3 中计算过 (回忆在定理 30.2 中我们曾得到在三维向量场 ξ_α 中的对分量 $\widehat{\Gamma}_{\beta\alpha\gamma}$ 的公式). 这时联络形式为

$$(A_\alpha)_{\beta\gamma} = \widehat{\Gamma}_{\beta\alpha\gamma}.$$

从公式 (30.1,10) 推导出, 当 α 为任意时 A_α 为反称的. 当群 G 由保持内积 $\langle \eta_1, \eta_2 \rangle = \delta_{\alpha\beta}\eta_1^\alpha \eta_2^\beta, \eta_i = \eta_i^\alpha \xi_\alpha, i = 1, 2$ 不变的变换组成; 因此, 群 G 为正交群, 并且度规变换在向量 η 上的作用由

$$\eta(x) \to g(x)\eta(x), \quad \eta = \eta^\alpha \xi_\alpha$$

给出, 其中 $g(x)$ 对每个 x 为正交的 $n \times n$ 矩阵.

因为像我们曾经注意到的那样, 联络 $\widehat{\Gamma}_{\beta\alpha\gamma}$ 对 β 和 γ 为反对称, 故它在正交群 G 的李代数中. 曲率形式 $(F_{\mu\nu})_{\kappa\lambda} = R_{\kappa\lambda,\mu\nu}$ 也在反称矩阵的李代数中取值:

$$R_{\kappa\lambda,\mu\nu} = -R_{\lambda\kappa,\mu\nu},$$

(见 §30). 对于伪黎曼度量一切都是类似的

对于在二维曲面上的度量情形, 曲率形式为

$$\Omega - K\sqrt{g}dx^1 \wedge dx^2, \tag{26}$$

其中 K 为曲面的高斯曲率.

c) 嘉当联络. G 为仿射群 (即线性变换和平移). 设在切向量上给出了任意的线性联络 $\nabla_\mu\xi^\nu$. 我们以公式

$$\widetilde{\nabla}_\mu\xi^\nu = \nabla_\mu\xi^\nu + \delta_\mu^\nu \tag{27}$$

定义"嘉当联络".

习题 41.4. 1) 证明联络 $\widetilde{\nabla}_\mu$ 对于形如

$$\xi \to g\xi + y \tag{28}$$

的局部仿射变换不变, 其中 $g = \left(\dfrac{\partial x^{\alpha'}}{\partial x^\alpha}\right)$ 为坐标变换 $x' = x'(x)$ 的矩阵, 而 $y = (y^1(x), \cdots, y^n(x))$ 为任意向量.

2) 证明这个联络的曲率形式 (取值于仿射群的李代数, 参看习题 4.3 和习题 24.11) 为

$$F_{\mu\nu} = (R^\kappa_{\lambda,\mu\nu}, T^\lambda_{\mu\nu}), \tag{29}$$

这里 $R^\kappa_{\lambda,\mu\nu}$ 为联络的曲率张量, $T^\lambda_{\mu\nu}$ 为挠率张量.

d) 旋量的共变微分以及度量

设 M^4 为时空空间, 度量为 $g_{\alpha\beta}$((1,3) 型). 与此度量相容的联络产生了旋量的共变微分. 我们将仔细分析半旋量表示的情形 (参看 §40.3). 这个表示的显式表达是这样的: 我们表示任意埃尔米特 2×2 矩阵 $U, \bar{U}^{\mathrm{T}} = U$ 的形式为

$$U = \begin{pmatrix} u^0 + u^3 & u^1 + iu^2 \\ u^1 - iu^2 & u^0 - u^3 \end{pmatrix} = u^\alpha\sigma_\alpha, \quad \bar{U}^{\mathrm{T}} - U, \tag{30}$$

其中 $\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \sigma_1, \sigma_2, \sigma_3$ 为泡利矩阵 (参看 §40); 矩阵 $\sigma_0, \cdots, \sigma_3$ 构成了 (实的) 埃尔米特矩阵的空间中的基. 于是矩阵 U 的行列式有形式

$$\det U = (u^0)^2 - (u^1)^2 - (u^2)^2 - (u^3)^2.$$

这个行列式在形如 $U \mapsto gU\bar{g}^{\mathrm{T}}$ 的变换保持不变, 其中矩阵 g 在群 $SL(2,\mathbb{C})$ 中. 因此, 变换 $U \mapsto gU\bar{g}^{\mathrm{T}}$ 可以看作是闵可夫斯基空间 (具坐标 u^0, u^1, u^2, u^3) 中保持度量不变的变换, 即群 $SO(3,1)$ 中的元. 我们得到对应 (同态) $SL(2,\mathbb{C}) \to SO(3,1)$, 这时矩阵 g 和 $-g$ 变成一个元. 逆映射 $SO(3,1) \to SL(2,\mathbb{C})$ 为双值, 这是个半旋表示.

作为群 $SL(2,\mathbb{C})$ 所作用的空间 $\mathbb{C}^2$ 中的元被称做 (两分量的) 旋量；称在这个空间中取值的场为旋量场. 在群 $SL(2,\mathbb{C})$ 的作用下, 它们按规律

$$\xi(x) \to g(x)\xi(x), \quad g(x) \in SL(2,\mathbb{C}) \tag{31}$$

变化. 也确定了共轭表示 $\xi \to \bar{g}\xi$.

按定义, 旋量场的共变导数为

$$\widehat{\nabla}_\alpha \xi = \frac{\partial \xi}{\partial x^\alpha} + A_\alpha \xi, \tag{32}$$

其中 A_α 为复的无迹 2×2 矩阵 (属于群 $SL(2,\mathbb{C})$ 的李代数). 我们要求对共轭旋量微分应按规律

$$\widehat{\nabla}_\alpha \xi = \frac{\partial \xi}{\partial x^\alpha} + \bar{A}_\alpha \xi \tag{32'}$$

进行.

对埃尔米特矩阵 $U(x)$ 的共变导数同样地定义 (莱布尼茨法则):

$$\widehat{\nabla} U = \frac{\partial U}{\partial x^\alpha} + A_\alpha U + U \bar{A}_\alpha^{\mathrm{T}}. \tag{33}$$

埃尔米特矩阵 U 利用等式 $U = u^\alpha \sigma_\alpha$ 与切向量 u^α 等同, 其中 (u^α) 是在向量 $e_0, e_1, e_2, e_3, e_0^2 = -e_1^2 = -e_2^2 = -e_3^2 = 1$ 的 "四元组" 中的切向量分量, 而向量 e_α 两两正交. 在这样的等同之下, 形如 (31) 的局部旋量变换对应于群 $SO(3,1)$ 中的局部洛伦兹变换. 我们再加上条件, 即由旋量联络 (33) 定义的共变导数 $\widehat{\nabla}_\alpha U$ 等同于表达式 $\sigma_\beta \nabla_\alpha u^\beta$, 其中 $\nabla_\alpha u^\beta$ 为向量 u^β 相对于与度量 $g_{\mu\nu}$ 相容的对称联络的共变导数:

$$\widehat{\nabla}_\alpha U = \sigma_\beta \nabla_\alpha u^\beta, \quad U = u^\beta \sigma_\beta \tag{34}$$

(这个联络在四元组中的分量已在 §30 中计算过).

习题 41.5. 证明, 旋量场的共变导数由条件 (32)—(34) 唯一地确定, 另外成立公式

$$A_\alpha = -\frac{1}{2}\widehat{\Gamma}^\gamma_{\beta\alpha}\sigma_\gamma\sigma_\beta \tag{35}$$

(对指标 β 是带符号的和, 即对应 $\beta = 0$ 的项取正号, 其余的取负号).

群 $SO(3,1)$ 的完全旋量表示分解为 $SL(2,\mathbb{C})$ 的表示和其复共轭的直和. 因此, 等式 (32)—(34) 允许共变微分也是四分量的旋量.

习题 41.6. 推导出度量存在时的狄拉克方程的形式.

注 每个旋量 $\xi = (\xi^0, \xi^1)$ 对应于埃尔米特矩阵 $U = (\xi^\alpha \bar{\xi}^\beta)$. 因为矩阵 U 为退化 (有零行列式), 故对应于它的向量 $u^k = \sigma^k_{\alpha\beta}\xi^\alpha\bar{\xi}^\beta$ 为迷向的.

习题 41.7. 我们选择旋量空间 $\mathbb{C}^2$ 的基 $\xi = (\xi^0, \xi^1), \eta = (\eta^0, \eta^1)$ 使得 $\xi^0\eta^1 - \xi^1\eta^0 = 1$. 证明此基典则地关联于闵可夫斯基空间 (四元组) 的标架, 在其中, 度量有

形式

$$(g_{ik}) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

§42. 度规不变的泛函的例子. 麦克斯韦和杨 – 米尔斯方程. 具恒等于零的变分导数的泛函 (示性类)

规范场的拉格朗日 $L = L\left(A_\mu, \frac{\partial A_\mu}{\partial x^\nu}\right)$ 应该具有下列性质:

1) $L\left(A_\mu, \frac{\partial A_\mu}{\partial x^\nu}\right)$ 为标量,

2) $L\left(A_\mu, \frac{\partial A_\mu}{\partial x^\nu}\right)$ 在度规变换下不变.

最简单的这种拉格朗日的形式为

$$L = -\frac{1}{4} g^{\mu\lambda} g^{\nu\kappa} \langle F_{\mu\nu}, F_{\lambda\kappa} \rangle, \tag{1}$$

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu} + [A_\mu, A_\nu].$$

这里 $g_{\mu\nu} = g_{\mu\nu}(x)$ 为所考虑的空间中的任意度量 : 以 $\langle , \rangle$ 表示群 G 的李代数的基灵形式

$$\langle X, Y \rangle = -\mathrm{Tr}(\mathrm{ad}\, X \mathrm{ad}\, Y),$$

其中 $\mathrm{ad}\, X(A) = [X, A]$(参照习题 24.9).(我们认为基灵形式非退化).

所构造的拉格朗日是个标量, 这是因为它是用在张量上的代数运算构成的. 我们来验证它的度规不变性. 在度规变换 (41.13) 下曲率形式 $F_{\mu\nu}$ 转换为 $gF_{\mu\nu}g^{-1}$. 由基灵形式的不变性 (见 §24) 我们有

$$\langle gF_{\mu\nu}g^{-1}, gF_{\lambda\kappa}g^{-1} \rangle = \langle F_{\mu\nu}, F_{\lambda\kappa} \rangle,$$

由此我们得到拉格朗日 (1) 的度规不变性.

设度量 $g_{\mu\nu}$ 为欧氏的或伪欧氏的:$g_{\mu\nu} = \varepsilon_\mu \delta_{\mu\nu}, \varepsilon_\mu = \pm 1$. 我们推导出泛函

$$S[A_\mu] = \int -\frac{1}{4} \langle F_{\mu\nu}, F_{\mu\nu} \rangle d^n x \tag{2}$$

极值的欧拉 – 拉格朗日方程 (对二重指标 μ, ν 求和, 并考虑进符号 $\varepsilon_\nu, \varepsilon_\mu = \pm 1$).

定理 42.1. 对泛函 (2) 的极值有方程

$$\nabla_\mu F_{\mu\nu} = 0, \tag{3}$$

其中 $\nabla_\mu F_{\mu\nu} \equiv \dfrac{\partial F_{\mu\nu}}{\partial x^\mu} + [A_\mu, F_{\mu\nu}]$ (参看 §41).

证明 对小局部变差 δA_μ 我们有

$$\delta S = -\frac{1}{2}\int \langle F_{\mu\nu}, \delta F_{\mu\nu}\rangle d^n x,$$

其中 $\delta F_{\mu\nu} = \dfrac{\partial}{\partial x^\mu}\delta A_\nu - \dfrac{\partial}{\partial x^\nu}\delta A_\mu + [\delta A_\mu, A_\nu] + [A_\mu, \delta A_\nu]$. 在 δS 的表达式中进行分部积分

$$\int \left\langle F_{\mu\nu}, \frac{\partial}{\partial x^\mu}\delta A_\nu \right\rangle d^n x = -\int \left\langle \frac{\partial F_{\mu\nu}}{\partial x^\mu}, \delta A_\nu \right\rangle d^n x,$$

并利用基灵形式的不变性

$$\langle F_{\mu\nu}, [\delta A_\mu, A_\nu]\rangle = -\langle [F_{\mu\nu}, A_\nu], \delta A_\mu\rangle,$$

我们得到

$$\delta S = \frac{1}{2}\int \left\{ \left\langle \frac{\partial F_{\mu\nu}}{\partial x^\mu}, \delta A_\nu \right\rangle - \left\langle \frac{\partial F_{\mu\nu}}{\partial x^\nu}, \delta A_\mu \right\rangle + \langle [F_{\mu\nu}, A_\nu], \delta A_\mu\rangle - \langle [F_{\mu\nu}, A_\mu], \delta A_\nu\rangle \right\} d^n x.$$

在指标的变换后我们得到

$$\delta S = \int \left\langle \frac{\partial F_{\mu\nu}}{\partial x^\mu} + [A_\mu, F_{\mu\nu}], \delta A_\nu \right\rangle d^n x = \int \langle \nabla_\mu F_{\mu\nu}, \delta A_\nu\rangle d^n x,$$

这时 δS 在任意变差 δA_ν 下, 在极值曲面上等于零. 故而对泛函 (2) 的极值我们得到了方程 $\nabla_\mu F_{\mu\nu} = 0$. 断言得证. □

注 这个方程 $\nabla_\mu F_{\mu\nu} = 0$ 和比安基恒等式 $\nabla_\lambda F_{\mu\nu} + \nabla_\mu F_{\nu\lambda} + \nabla_\nu F_{\lambda\mu} = 0$ 在没有基灵度量的非退化的假定时也可以被描述出来. 但是它们将没有 (在所有简单的情形中) 拉格朗日形式.

习题 42.1. 对于仿射群的嘉当联络, 在度量和曲率之间的关系置换之后的这些方程具有爱因斯坦方程的形式 (由现代的研究工作提取出来的)

$$R_{ab} - \frac{1}{2}R g_{ab} = 0. \tag{4}$$

例 42.1. 在阿贝尔情形 ($G = U(1)$ 为电磁场) 我们得到了拉格朗日

$$L = -\frac{1}{4}\left(\frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}\right)^2 = -\frac{1}{4}F_{\mu\nu}^2, \tag{5}$$

即电磁场的标准拉格朗日 (参看 §37.3). 方程 (3) 在这里等同于麦克斯韦方程

$$\frac{\partial F_{\mu\nu}}{\partial x^\mu} = 0. \tag{6}$$

注　对于群 $G = SU(2)$, 规范场 A_μ 被常称做杨－米尔斯场.

习题 42.2.　推导拉格朗日

$$L = -\frac{1}{4} g^{\mu\lambda} g^{\nu\kappa} \langle F_{\mu\nu}, F_{\lambda\kappa} \rangle$$

的极值方程 $\dfrac{\delta S}{\delta A} = 0$, 其中度量 $g_{\mu\nu}(x)$ 为任意而固定 (外部场) 的.

在四维空间中形如

$$S[A] = \int \mathrm{Tr}\ F_{\mu\nu}^2 d^4 x$$

的作用有一个附加的对称性, 即在 $\mathbb{R}^4$(或在 $\mathbb{R}^4_{3,1}$) 中所有共形变换的群下的不变性 (参看 §15 中对此群的描述). 对任意共形变换的微分则化作拉伸与旋转 (见 §15), 故而只要对于拉伸 $x \to \lambda x$ 验证共形不变即可. 在此变换下我们有

$$F_{\mu\nu} \to \lambda^{-2} F_{\mu\nu}\ (F_{\mu\nu}\text{为张量}), \quad d^4 x \to \lambda^4 d^4 x,$$

其中 d^4x 为体积元. 由此显然有

$$\mathrm{Tr}\ F_{\mu\nu}^2 d^4 x \to \mathrm{Tr}\ F_{\mu\nu}^2 d^4 x.$$

结论　拉格朗日 (2) 和在闵可夫斯空间中的麦克斯韦－杨－米尔斯方程为共形不变, 即具有对称群 $O(4,2)$(见 §15).

度规不变的微分形式也起着重要的作用, 这里的微分形式具有依赖于 A_μ 的标量值. 例如二阶形式

$$c_1 = \mathrm{Tr}\ \Omega = \sum_{\mu<\nu} \mathrm{Tr}\ F_{\mu\nu} dx^\mu \wedge dx^\nu \tag{7}$$

为度规不变

$$\mathrm{Tr}(g\Omega g^{-1}) = \mathrm{Tr}\ \Omega.$$

这里的 Tr 表示矩阵的迹. 形式 $c_1 = \mathrm{Tr}\ \Omega$(局部地) 为恰当微分 (全微分):$\mathrm{Tr}\ \Omega = d(\mathrm{Tr}\ A), A = A_\mu dx^\mu$. 事实上, 由于矩阵的换位子的迹为零, 故

$$\begin{aligned}\mathrm{Tr}\ \Omega &= \sum_{\mu<\nu} \mathrm{Tr}\left(\frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu} + [A_\mu, A_\nu]\right) dx^\mu \wedge dx^\nu \\ &= \sum_{\mu<\nu} \mathrm{Tr}\left(\frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}\right) dx^\mu \wedge dx^\nu = d\mathrm{Tr}\ A.\end{aligned}$$

联络的局部变差 $A_\mu \to A_\mu + \delta A_\mu$ 下, 形式 $\mathrm{Tr}\ \Omega$ 变为形式 $\mathrm{Tr}\ \Omega + \mathrm{Tr}\ \delta\Omega$, 其中 $\mathrm{Tr}\ \delta\Omega = d\mathrm{Tr}\ \delta A$ 为恰当形式. 因此泛函

$$S_1[A_\mu] = \int_{\mathbb{R}^2} \mathrm{Tr}\ \Omega$$

(在二维空间上) 由于变分的局部性, 故有恒等于零的变分导数

$$\delta S_1[A_\mu]=\int_{\mathbb{R}^2} d(\mathrm{Tr}\ \delta A)=0, \tag{8}$$

定义 42.1. 闭的度规不变形式 ω 使泛函 $\int\omega$ 具有恒等于 0 的变分导数 $\frac{\delta\int\omega}{\delta A}=0$, 则被称为示性类 (微分几何的).

另外的示性类为

$$c_i=\mathrm{Tr}(\Omega\wedge\cdots\wedge\Omega)=\mathrm{Tr}(\Omega^i),\quad i\geqslant 1. \tag{9}$$

c_i 为 $2i$ 阶的度规不变形式;(9) 中形式 Ω 的分量像矩阵那样做乘法.

习题 42.3. a) 证明当 $i\geqslant 1$ 时形式 c_i 为闭.

b) 证明泛函

$$S_i[A_\mu]=\int_{\mathbb{R}^{2i}} c_i$$

(在 $2i$ 维空间上) 有恒等于零的变分导数

$$\frac{\delta S_i[A]}{\delta A_\mu}\equiv 0.$$

c) 证明这也对形如

$$\int_{\mathbb{R}^4}(\alpha c_1\wedge c_1+\beta c_2),\quad \int_{\mathbb{R}^6}(\alpha c_1^3+\beta c_1\wedge c_2+\gamma c_3)$$

等等, 对任意多的 c_i 项也都对.

对于群 $G=SO(2n), c_1=c_3=c_5=\cdots=0$; 除了只在 $p_i=c_{2i}$ 的非平凡的情形外, 还有一个示性类

$$\chi_n=\varepsilon^{i_1\cdots i_{2n}}\Omega_{i_1i_2}\wedge\Omega_{i_2i_3}\wedge\cdots\wedge\Omega_{i_{2n-1}i_{2n}}, \tag{10}$$

其中 $\varepsilon^{i_1\cdots i_{2n}}$ 为置换 $\begin{pmatrix}1\cdots 2n\\ i_1\cdots i_{2n}\end{pmatrix}$ 的符号,Ω_{ij} 是曲率形式 $\Omega=\sum_{\mu<\nu}F_{\mu\nu}dx^\mu\wedge dx^\nu$ 的矩阵分量.

例 42.2. 当 $n=1$, 我们得到 χ_1 的表达式 (参看前一节)

$$\chi_1=\varepsilon^{i_1i_2}\Omega_{i_1i_2}=2K\sqrt{g}dx^1\wedge dx^2. \tag{11}$$

我们已知 (见定理 37.3) 泛函

$$S[g_{ij}]=\int_{\mathbb{R}^2}K\sqrt{g}dx^1\wedge dx^2-\int_{\mathbb{R}^2}\frac{1}{2}R\sqrt{g}dx^1\wedge dx^2 \tag{12}$$

有对于度量的恒等于零的变分导数 (参看 §37).

例 42.3. 当 $n=2$, 我们有 χ_2 和 c_2. 对于黎曼度量有

$$\chi_2 = \varepsilon^{ijkl} R_{ij\mu\nu} R_{kl\rho\sigma} dx^\mu \wedge dx^\nu \wedge dx^\rho \wedge dx^\sigma,$$

$$c_2 = R_{ij\mu\nu} R^{ij}_{\rho\sigma} dx^\mu \wedge dx^\nu \wedge dx^\rho \wedge x^\sigma,$$

习题

42.4. 证明所有形式 χ_n 为闭的, 并确为示性类.
对于 $G=SO(2n)$ 的一般的示性类为 $\chi_n, c_2, c_4, \cdots, c_{2n-2}$ 的多项式.

42.5. 证明对 $G=SO(4)$, 泛函

$$\int\limits_{\mathbb{R}^4} \chi_2, \int\limits_{\mathbb{R}^4} c_2, \int\limits_{\mathbb{R}^8} \chi_2^2, \int\limits_{\mathbb{R}^8} \chi_2 c_2, \int\limits_{\mathbb{R}^8} c_2^2$$

有恒为零的变分导数, 即所有形式 $\chi_2, c_2, \chi_2^2, \chi_2 c_2, c_2^2$ 都是示性类.

参考文献

Ⅰ. 几何和拓扑方面的教科书

1. *Рашевский П. К.* Риманова геометрия и тензорный анализ. —М. :Наука,1967.
2. *Рашевский П. К.* Курс дифференциальной геометрии. —М. ;Гостехиздат,1956.
3. *Погорелов А. В.* Дифференциальная геометрия. —М. :Наука,1974.
4. *Погорелов А. В.* Внешняя геометрия выпуклых поверхностей. —М. ; Наука, 1969.
5. *Александров А. Д.* Внутренняя геометрия выпуклых поверхностей. —М. ; Л. : Гостехиздат,1948.
6. *Ефимов Н. В.* Высшая геометрия. —М. : Наука,1971.
7. *Норден А. П.* Теория поверхностей. —М. ;Гостехиздат,1956.
8. *Фиников С. П.* Курс дифференциальной геометрии. —М. ;Гостехиздат,1952.
9. Seifert, H., Threlfall, W., Topoloqie, 1934(有中译本).
10. Seifert, H., Threlfall, W., Variationsrechnung in Crossen. Hamburger Math Einzelschr., No.24,Taubner;Leipzig.
11. Milnor, J. W., Morse Theory. Ann. of Math. Studies, No. 51. Princeton Univ. Press;Princeton, N. J,1963.
12. Milnor, J. W., Singular Points of Complex Hypersurfaces. Ann. of Math. Studies, No.61, Princeton Univ. Press: Princeton, N.J.
13. *Понтрягин Л. С.* Гладкие многообразия и их применения в теории гомотопий. —М. ;Наука,1985.
14. *Понтрягин Л. С.* Непрерывные группы. —М. : Наука,1984.(有中译本)
15. Serre, J. P., Lie Algebras and Lie Groups. Lectures given at Harvard Univ. W. A. Benjamin: New York. 1964.
16. Springer, G., Introduction to Riemann Surfaces. Addison – Wesley;Reading, Mass. 1957.

17. Nomizu, K., Lie Groups and Differential Geometry. Math. Soc. Japan. 1956.

18. Chern, S. S., Complex Manifolds. Textos de Matematica, No. 5. Instituto do Fisica e Matematica. Universidado do Recife, 1959.

19. Bishap, R. L. and Czittenden, R. J., Geometry of Manifolds. Pure and Applied Math. Vol. 25. Academic Press: New York – London, 1964.

20. Gromoll, D., Klingenberg, W., and Meyer, W., Riemannsche Geometrie in Grossen. LNM No. 55, Springer – Verlag: Berlin – New York. 1968.

21. Helgason, S., Differential Geometry and Symmetric Spaces. Academic Press: New York. 1968.

22. Steenrod, N. E., The Topology of Fibre Bundles. Princeton Math. Series, Vol. 14, Princeton Univ. Press: Princeton, N. J.

23. *Розендорн Э. Р.* Задачи по дифференциальной геометрии. —М.:Наука,1971.

24. *Новиков С. П., Мищенко А. С., Соловьев Ю. П., Фоменко А. Т.* Задачи по геометрии. —М.:Изд-во Моск. ун-та,1978.

25. Hilbert, D. and Cohn – Vossen, S., Geometry and the Imagination. Chelsea Publishing Co.: New York(有中、英译本). 1952.

26. *Рохлин В. А., Фукс Д. Б.* Начальный курс топологии. Геометрические главы. —М.:Наука,1977.

27. Lefschetz, S., Algebraic Topology. Am. Math. Soc. Colloquium Publications, Vol. 27. 1942.

28. *Дубровин Б. А., Новиков С. П., Фоменко А. Т.* Современная геометрия. Методы и приложения. Т. 3: Теория гомологий. —М.: Эдиториал УРСС, Добросвет,2001.(即本书第三卷).

Ⅱ. 关于微分方程和经典力学方面的书

29. *Понтрягин Л. С.* Обыкновенные дифференциальные уравнения. —М.: Наука,1969.(有中译本)

30. *Арнольд В. И.* Обыкновенные дифференциальные уравнения. —М.:Наука, 1974.

31. *Арнольд В. И.* Дополнительные главы теории обыкновенных дифференциальных уравнений. —М.:Наука,1978.

32. *Арнольд В. И.* Математические методы классической механики. —М.: Эдиториал УРСС,2000.(有英译本)

33. Coddington, E. A. and Levinson, N., Theory of Ordinary Differential Equations. McGraw – Hill: New York – Toronto – London, 1955.

34. *Ландау Л. Д., Лифшиц Е. М.* Механика. —М.:Наука,1973.(有英译本)

35. *Голубев В. В.* Лекции по интегрированию уравнений движения тяжелого твердого тела около неподвижной точки. —М. :Гостехиздат,1954.

Ⅲ. 补充文献

36. *Ахиезер А. И.*, *Берестецкий В. Б.* Квантовая электродинамика. —М. :Наука, 1969.

37. *Боголюбов Н. Н.*, *Ширков Д. В.* Введение в теорию квантованных полей. —М. :Наука,1984.(有英译本)

38. *Ландау Л. Д.*, *Лифшиц Е. М.* Теория поля. —М. :Наука,1973.(有英译本)

39. Misner, R. E., Thorne, K. S. and Wheeler, J. A., Gravitation. W. H. Freeman: San Francisco, 1973.

40. *Седов Л. И.* Механика сплошной среды. —М. :Наука, 1983,1984.

41. *Ландау Л. Д.*, *Лифшиц Е. М.* Механика сплошных сред. —М. :Гостехиздат, 1954.(有英译本)

42. Peieris, R. E., Qumtum Theory of Solids. Theoretical Physics in the Twentieth Center. Interscience: N. Y. 1960.

43. Feynman, R. P., Leighton, R. B. and Sands, M., The Feynman Lectures on Physics. Addison – Wesley: Reading, Mass., 1963.

44. Wintner, A., Analytical Foundations of Celestial Mechanics. Princeton Univ. Press: Princeton, N. J. 1941.

45. *Славнов А. А.*, *Фаддеев Л. Д.* Введение в квантовую теорию калибровочных полей. —М. : Наука, 1978.

46. *Делоне Б. Н.*, *Александров А. Д.*, *Падуров Н. Н.* Математические основы структурного анализа кристаллов. —Л. ;М. :ОНТИ,1934.

47. *Зельдович Я. Б.*, *Новиков И. Д.* Строение и эволюция Вселенной. —М. : Наука,1975.

48. Теория солитонов/Под ред. С. П. Новикова. —М. : Наука,1979.

49. *Богоявленский О. И.* Методы качественной теории динамических систем в астрофизике и газовой динамике. —М. :Наука, 1980.

50. Нелинейные волны. —М. :Наука, 1979,с.60 –72.

索 引

A
埃尔米特 (Hermite) 度量 208
埃尔米特内积 73
爱因斯坦 (Einstein) 方程 302
爱因斯坦假设 322

B
半旋表示 335
保有电流 316
贝尔特拉米 (Beltrami) 方程 85
贝尔特拉米算子 85
彼得松 – 柯达西 (Peterson-Codazzi) 方程 241
变分导数 247,291
波前 284
泊松 (Poisson) 括号 271
布喇菲 (Bravais) 格 136

C
残数公式 200
测地线 223
超曲面 50
超椭圆曲线 82

D
带边界的区域 3
单参数子群 98
等温坐标 83
狄拉克 (Dirac) 方程 336
狄拉克共轭 335
狄利克雷 (Dirichlet) 泛函 310
笛卡儿 (Descartes) 坐标 1
典则变换 276
典则变换的生成函数 279
电磁波 155
电磁场 149
电磁场的不变量 149
电磁场的势向量 266
电磁场的作用 298
电磁场能量 – 动量张量 153,301
定向曲线的曲率 31
动量 247
动量矩 257
动量矩张量 296
度规变换 340
对称晶体群 140
对称联络 219
对称线性算子 132
多向量 128

E

二次型 (在向量上的)　13
二次型的特征值　132

F

发散量　182
反称线性算子　132
反称张量的积分　192
反称张量的梯度　181
反交换变量的积分　129
反射旋转　26
反演　104
仿射群　21
非异拉格朗日　261
非异曲面　48
费马 (Fermat) 原理　267
弗莱纳 (Frenét) 公式　32
复标量场拉格朗日　316
复分析函数　78
复空间区域的变换　80
复解析坐标变换　79
复向量场的拉格朗日　317
复雅可比　79

G

高斯 (Gauss) 曲率　59,62
高斯映射　196
格的点群　140
格的基本胞腔　136
格的空间群　138
共变常向量场　222
共变微分　217
共轭点　288
共形度量　103
共形变换　23
共形坐标　83
固有时间　41
规范场　340
规范群　340

H

哈密顿 (Hamilton)　269
哈密顿方程　261
哈密顿体系　269
哈密顿 – 雅可比 (Jacobi) 方程　280
含电磁场　265,337
函数在曲面的积分　193
黑塞 (Hesse)　59
胡克 (Hooke) 定律　119
滑动反射　22

J

基灵 (Killing) 度量　172
积分曲线　157
极小曲面　308
极坐标　5
极值曲线　246
嘉当 (Cartan) 联络　344
伽利略 (Galileo) 群　29
降标　131
截尾方程　282
晶体格　135
晶体平移群　138
矩阵的指数映射　96

K

凯莱 (Kähler) 度量　208
柯西 – 黎曼 (Cauchy-Riemann) 方程　78
克莱因 (Klein) 模型中的度量　88
克莱因 – 戈登 (Jordon) 方程　316
克里斯托费尔 (Christoffel) 公式　225
克氏符号　217
空间的维数　2

L

拉格朗日 (Lagrange)　251
拉格朗日的曲面　280
拉格朗日圆锥曲面　283
拉普拉斯 – 龙格 – 楞茨 (Laplace-Runge-Lenz)积分　274
勒让德 (Legendre) 变换　261
类空向量　40

类时向量 40
黎曼 (Riemann) 面 82
黎曼度量 13
黎曼曲率 231
黎曼张量 231
李 (Lie) 代数 164
李导数 160
里奇 (Ricci) 张量 236
力 247
连续介质能量 – 动量张量 328
刘维尔 (Liouville) 方程 90
罗巴切夫斯基 (Lobachevski) 度量 69
洛伦兹 (Lorenz) 变换 41,44
洛伦兹群的旋表示 333

M

麦克斯韦 (Maxwell) 方程 188
慢的测地线 255
迷向 (光) 锥 40
迷向向量 40
莫佩尔蒂 (Maupertuis) 原理 266
闵可夫斯基 (Minkowski) 空间 17

N

挠率张量 219
内积 14,115
能量 246
能量 – 张量向量 254
能量 – 动量张量 294

O

欧拉 (Euler) 公式 63
欧拉 – 拉格朗日方程 247
欧拉角 94
欧氏度量 15
欧氏空间 6
欧氏内积 7
欧氏坐标 6

P

庞加莱 (Poincaré) 模型的度量 88
庞加莱群 41
泡利 (Pauli) 矩阵 99
平均曲率 59,62
普法夫 (Pfaff) 269

Q

强非异拉格朗日 261
球极投影 67,70
球面度量 66
球面坐标 6
区域的边界 3
区域的变换 18
曲率半径 31,34
曲率形式 196,341
曲面的第二基本型 59
曲面的第一基本型 52
曲面的非异点 48
曲面上的度量 52
曲面上的局部坐标 49
曲面上区域的面积 55
曲线的长度 14
曲线的非异点 36
曲线的挠率 36
曲线的切向量 7
曲线的曲率 31,34
曲线的速度向量 7,10
曲线的自然方程 37
曲线间的夹角 8
全纯形式 208
群 18

R

弱场 254

S

升标 132
施瓦氏 (Schwarzchild) 度量 326
时空间隔 17
时空连续体 2
时空平面 301

实标量场拉格朗日 317
实化 72
示性类 349
数值曲率 236
斯托克斯 (Stokes) 公式 199
四元数 99
四元体系 233
缩并 121

T

梯度系统 268
调和径向量 309
通过曲面的张量场的流 196

W

外代数 128
外微分 183
外尔 (Weyl) 方程 337
微分几何的联络 218
微分形式 123
伪黎曼度量 16
伪欧氏度量 16
伪欧氏坐标 16
伪球面坐标 17
无边界区域 2

X

线性算子 115
线性算子的迹 121
线性向量场 170
相对论的极小作用原理 252
相空间 261,269
向量 11
向量场的换位子 162
向量场的积分 161
向量场的指数映射 159
向量的平移 222
向量积 35
向量与余向量的内积 123
小形变张量 119
辛变换 277
形变 157
形变张量 118
旋度 182
旋转群的旋表示 333
循环向量场 195

Y

雅可比 (Jacobi) 4
雅可比恒等式 164
雅可比场 288
雅可比算子 286
杨 – 米尔斯 (Yang-Mills) 场 348
引力半径 326
引力场 254
引力场的作用 301
应力张量 118
映射的微分 157
酉变换 74
酉群 74
余向量 12
余向量的内积 115
与度量相容的联络 224
运动 19
运动群 19

Z

张量 116
张量的乘积 122
张量的限制 156
正常洛伦兹变换 43
正常运动 22
正时序变换 43
正弦 – 戈登 (Sine-Gordon) 方程 242
指标置换 121
质点的世界线 2
主法线 36
主曲率 59
主曲率方向 61
柱面坐标 5

主曲率方向　61
柱面坐标　5
自然参数　10
最小作用原理　247
左不变向量场　172
作用　244
坐标变换的雅可比矩阵　4
坐标系的非异点　4
4 元向量流　298
n 维笛卡儿空间　1
γ 矩阵　333
$*$ 算子　133
(p, q) 型形式　127